

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2002, 13:36:02 ; Search time 157.61 Seconds
(without alignments)
7900.362 Million cell updates/sec

Title: US-09-715-909-1

Perfect score: 5498
Sequence: 1-cataatacacaataagagga.....aaaaaaaaaacctgcag 5498

Scoring table: IDENTITY_NUC
Gap 10.0 , Gapext 1.0

Number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- 1: /cg2.6/ptodata/2/ina/5A.COMB.seq:*
- 2: /cg2.6/ptodata/2/ina/5B.COMB.seq:*
- 3: /cg2.6/ptodata/2/ina/6A.COMB.seq:*
- 4: /cg2.6/ptodata/2/ina/6B.COMB.seq:*
- 5: /cg2.6/ptodata/2/ina/PCBUS.COMB.seq:*
- 6: /cg2.6/ptodata/2/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1954.6	35.6	5577	1	US-08-326-117B-1
2	1954.6	35.6	5577	3	US-08-982-129-1
3	57.2	1.0	790	6	US-08-982-129-1
4	55	1.0	1910	5	PCT-US92-05532-1
5	55	1.0	2261	1	US-08-272-882D-1
6	53.6	1.0	789	6	US-08-272-882D-1
7	53.2	1.0	1294	3	US-09-025-691-2
8	51	0.9	3377	6	US-08-453-16
9	50.8	0.9	4104	1	US-07-998-003A-94
10	50.8	0.9	4104	1	US-08-453-274B-94
11	50.8	0.9	4104	1	US-08-453-695A-94
12	50.8	0.9	4104	1	US-08-268-161A-94
13	50.8	0.9	4104	2	US-08-453-702A-94
14	50.8	0.9	4104	4	US-09-099-639-94
15	50.8	0.9	4104	5	PCT-US93-12588-94
16	50.8	0.9	4104	5	PCT-US95-08071-94
17	50.8	0.9	4650	1	US-07-998-003A-102
18	50.8	0.9	4650	1	US-08-453-274B-102
19	50.8	0.9	4650	1	US-08-453-695A-102
20	50.8	0.9	4650	1	US-08-268-161A-102
21	50.8	0.9	4650	2	US-08-453-702A-102
22	50.8	0.9	4650	2	US-09-099-639-102
23	50.8	0.9	4650	5	PCT-US93-12588-102
24	50.8	0.9	4650	5	PCT-US95-08071-102
25	49.2	0.9	1288	1	US-08-440-856A-9
26	49.2	0.9	4972	3	US-09-035-648-17
27	49.2	0.9	4972	4	US-09-001-951-17

28	48.4	0.9	1302	1	US-08-525-697-1	Sequence 1, Appl
29	46.8	0.9	1656	1	US-08-741-881-1	Sequence 1, Appl
30	46.8	0.9	1656	1	US-08-739-158-1	Sequence 1, Appl
31	46.8	0.9	1656	2	US-08-739-167-1	Sequence 1, Appl
32	46.8	0.9	1656	3	US-08-404-796-1	Sequence 1, Appl
33	46.8	0.9	1656	3	US-08-931-869-1	Sequence 1, Appl
34	46	0.8	1272	2	US-09-068-109-1	Sequence 1, Appl
35	46	0.8	2712	3	US-09-025-691-4	Sequence 1, Appl
36	45.8	0.8	2007	3	US-08-747-221B-36	Sequence 36, Appl
37	45.8	0.8	2007	3	US-08-747-221B-38	Sequence 36, Appl
38	45.8	0.8	2007	4	US-09-005-051-36	Sequence 38, Appl
39	45.8	0.8	2007	4	US-09-005-051-38	Sequence 38, Appl
40	44.6	0.8	1035	4	US-09-105-390-49	Sequence 49, Appl
41	44.4	0.8	2230	4	US-09-105-390-1	Sequence 1, Appl
42	44.2	0.8	390	4	US-09-197-649-7	Sequence 7, Appl
43	44.2	0.8	924	4	US-09-105-390-33	Sequence 33, Appl
44	44	0.8	1431	4	US-09-078-173A-7	Sequence 7, Appl
45	44	0.8	1443	4	US-09-078-173A-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-08-326-117B-1
Sequence 1 Application US/08326117B
Patent No 5693491
GENERAL INFORMATION:
APPLICANT: BULLA, LEE A.
APPLICANT: JI, TAE
TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
TITLE OF INVENTION: TOXIN
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/326,117B
FILING DATE: 19-OCT-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MILMAN, ROBERT A
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 7112-0037.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5577 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 197..4780
US-08-326-117B-1

Query Match 35.6%, Score 1954.6; DB 1; Length 5577;
Best Local Similarity 63.9%; Pred. No. 0;
Matches 3248; Conservative 0; Mismatches 1769; Indels 62; Gaps 17;

Qy	152	atcaltgcaaaatg	ggg	gggtg	gtgag	aggtc	tctc	ccagc	agtg	tact	ggtc	tctt	ttg	ccctc	211
Db	187	ATATGTGAGATG	AGCATG	AGTGC	AGTCCG	AAATGCTG	CTGCTG	CTGCTG	CTGCTG	CTGCTG	CTGCTG	CTGCTG	CTGCTG	CTGCTG	246
Qy	212	tgccgcactag	ccaa	ccca	agatg	tctg	tact	ta	tcg	ta	tcg	ta	tcg	ta	271
Db	247	TGCAGTTT	TACCTC	AAGAGATG	TGGGTAT	TGACGG	CCATCC	CAAGGCT	TAC	CAAGGCT	TAC	CAAGGCT	TAC	CAAGGCT	306
Qy	272	ggaac---	tgccgc	actatg	ta	tcg	aa	gaa	aatc	ta	gag	ta	gag	ta	328
Db	307	GGATAAATTT	TGCGAG	TACTAAAT	TTTGAAG	GGCCAG	ACATG	AGTCA	GAGAG	GGCCCT	GTCTCC				366
Qy	329	cggcccgac	cccgag	gag	ag	ta	tgt	atg	aga	ac	-----	tcttc	cg	gata	379
Db	367	CGCCCGGAGG	GGG	GTGAC	CTGTG	ATGAG	CGCTT	ACATCC	AGTAA	CAACCA	CTCGG				426
Qy	380	gattcag	tcatat	cata	tcg	gag	aga	gaa	atc	ga	agag	ac	g	ta	439
Db	427	CACGCAAG	GTCACT	CTACAT	TGATG	TAAGAG	ATAGTA	AGACGA	AAATC	ACCATG	CGCCCT	TACTTAA			486
Qy	440	ctatca	agg	gtc	ca	ac	gc	gc	gtg	tc	g	ta	ta	ta	499
Db	487	TTAT-AA	CGAGCA	CTCACT	CCGTT	CATG	GAAC	TGCACT	GGCACT	TTTAT	CCGGT	CTGTAC	CAATC		545
Qy	500	gggccc	ta	gtt	tc	ga	cga	atg	aa	gc	ag	cc	ta	g	559
Db	546	TGCTGAT	CCCGG	TCAT	CAGAG	AGATTG	ACACG	GGGAG	GTGCAT	-TCAT	ATCAT	ACGCAAG			603
Qy	560	acaag	acta	cga	ag	ca	gc	ac	ca	tc	ga	gc	ta	g	619
Db	604	ACAGCATTA	ACGAGTT	GTCCCG	GATG	CAGAGTA	CAATGTT	CAATGTT	CAATGTT	CAATGTT	CAATGTT	CAATGTT	CAATGTT	CAATGTT	663
Qy	620	atcaca	agg	cc	gt	act	gtg	tc	g	aga	ta	g	ta	g	679
Db	664	GTCCGTG	GGGAG	CGGCTG	CTCTC	TGCTAT	CGTCA	ACATG	ATGATG	ATGATG	ATGATG	ATGATG	ATGATG	ATGATG	723
Qy	680	gcaag	tg	tca	gc	gc	tc	gt	ta	at	tc	ga	aa	tc	739
Db	724	ACAAAAC	TTTGA	GCCTT	GGCG	GTCTT	CTA	GAAC	TGGG	CGCC	AGG	GTG	AGCA	AAATG	783
Qy	740	gtacca	agt	gt	caga	cc	gc	gc	gc	gc	gc	gc	gc	gc	799
Db	784	ATACCA	AGAT	TGAG	ACGCGG	AGCGAG	CGGAT	ACACAC	AGATT	CATG	AGCTT	CATG	AGAT	CGA	843
Qy	800	caag	ca	gg	gc	tc	ga	ta	g	aa	gc	a	tc	t	859
Db	844	CAGCGT	TCG	---TG	CGAG	GAGAG	ACCTT	CTAT	CA	TGAA	CGGAC	GAAT	TCCCA	ACCA	900
Qy	860	cggt	tca	ata	ga	aga	ta	gc	ta	gc	ta	gc	ta	gc	919
Db	901	ATGGA	ATG	GGTAA	ATATG	ACATAG	GGCG	TTAA	TAC	CTG	CTCA	CTT	CGT	CA	960
Qy	920	gcttca	tata	t	ta	gt	c	ta	c	gc	t	c	ta	c	979
Db	961	GCTG	CTAT	TAT	TAC	AGG	GTAG	ACAC	CCCT	GGAC	CTG	CTG	CTG	CTG	1020
Qy	980	gat	tg	tca	ag	tg	aga	gc	ta	g	at	gc	tc	gc	1039
Db	1021	GATG	TG	CGA	AGTGA	AGTGA	ACAG	CGCT	CCG	CG	CG	CG	CG	CG	1080
Qy	1040	gaag	cg	gtt	ga	g	aa	ga	gc	ta	gc	ta	gc	ta	1099
Db	1081	CCAC	AGTT	TGA	GA	GA	AAAT	TTTAC	CA	AAAT	TTTAC	AGTA	GGCG	AT	1140
Qy	1100	ggga	at	c	a	a	t	a	a	g	c	a	a	t	1159
Db	1141	TGAG	AT	CA	AT	TAT	TCC	TA	CTA	CA	CTA	CA	CTA	CA	1200
Qy	1160	cctg	ag	gt	at	-----	ga	ag	cg	ag	ct	at	tc	tc	1210
Db	1201	CATT	G												

Qy	1261	CGACACAC	CTGC	AAAGAG	GTTC	ACACT	TAG	CTGT	GCCTT	ACAAAT	TGAT	GAG	CA	1320	
Qy	1271	ta	gctt	gca	a	g	a	g	g	c	c	a	a	t	1330
Db	1321	GGCTT	CCACAT	CAACAA	ACG	GTGAT	CAT	TGTG	CA	CA	GAT	CA	CA	1380	
Qy	1331	cgag	cg	gc	t	ca	a	a	a	a	a	a	a	a	1390
Db	1381	TGA	CC	TAT	TAC	ACAG	AAAT	TG	CAC	TG	CG	CA	TAT	G	1440
Qy	1391	cttca	a	ga	a	a	a	a	a	a	a	a	a	a	1450
Db	1441	CTTG	AT	TA	AA	GA	AT	TC	GA	AT	T	T	A	G	1500
Qy	1451	gga	act	t	ga	a	gc	gt	tc	cc	g	a	g	g	1510
Db	1501	GCGT	TA	GAG	AG	CGCT	GC	AC	CC	T	CA	AG	CG	CT	1560
Qy	1511	cg	gt	tca	ca	a	g	g	a	a	a	a	a	a	1570
Db	1561	CGG	CT	TAC	AG	CG	AC	AC	CTT	CAT	CAT	G	G	A	1620
Qy	1571	ag	at	g	t	c	a	t	t	t	t	t	t	t	1630
Db	1621	AGTG	CC	AG	AT	T	T	T	T	T	T	T	T	T	1680
Qy	1631	cc	a	g	t	g	g	g	a	g	c	g	c	t	1690
Db	1681	GC	AG	T	G	G	G	C	T	G	G	T	T	T	1740
Qy	1691	ctt	g	a	a	g	a	g	a	g	a	g	a	g	1750
Db	1741	CTTG	AC	AC	CC	CG	GT	GC	AG	AC	CG	T	T	T	1800
Qy	1751	g	c	a	a	g	t	g	t	c	g	c	c	t	1810
Db	1801	CG	CA	AG	CG	G	T	T	G	C	A	C	A	C	1860
Qy	1811	caa	g	c	t	t	a	c	c	t	t	a	c	c	1870
Db	1861	T	A	G	C	T	T	A	A	C	T	T	C	T	1920
Qy	1871	cg	a	g	c	c	t	t	a	c	c	a	a	c	1930
Db	1921	CG	AC	T	C	T	T	C	A	C	T	C	A	C	1980
Qy	1931	t	t	g	g	c	a	g	c	y	c	c	a	a	1990
Db	1981	G	T	G	G	G	C	A	---A	C	C	T	T	C	2030
Qy	1991	caa	a	a	a	c	t	c	c	c	c	c	c	c	2050
Db	2038	C	A	A	C	A	C	G	C	C	C	T	T	A	2090
Qy	2051	t	c	c	g	a	a	g	a	t	a	t	c	c	2110
Db	2098	G	C	T	A	T	G	C	A	C	A	T	C	A	2150
Qy	2111	c	c	t	g	t	t	c	g	a	t	c	a	t	2170
Db	2158	T	C	T	G	C	T	T	G	A	A	A	A	A	2210
Qy	2171	c	c	c	t	a	t	a	c	a	c	a	c	a	2230
Db	2218	C	C	C	C	A	C	A	G	T	T	A	G	A	2270
Qy	2231	g	g	t	t	c	c	a	c	c	g	c	c	t	2290
Db	2278	G	G	A	C	T	G	T	A	C	G	G	G	T	2330
Qy	2291	c	g	a	a	a	t	c	g	a	t	c	t	a	2350

[illegible]

Db	3415	GCTGGAGACCCGCATNGACCTCAAGGAGATTTGGGGAGCGTACGCTAATATACGAGGCG	3474
Qy	3422	cttcgacacagcgctacacccgcagcagaagctccaaacagagactcaaaccttgcataccgcgc	3481
Db	3475	ATTGCACGACGGCATTTCGGCAAAATGTCATTAAGAACGAGACATATAGAGCTATCATCATCC	3534
Qy	3482	ctcaaacctccacaccccttgttctgttccgtccgaacccgaactccgtcatctcgtcttc	3541
Db	3535	GTTCAACTACTACCGCGCTGATTCGCTCTCCGACCAACGATGCGTCATACGACTTGC	3594
Qy	3542	taggagagcgcaacagaaagcgcgctctggagagcgctgcacaagagttcttggagcc	3601
Db	3595	GAGGGAAAGACCTTATTCATATGAGATTTTAGGAGAAAGMAACGAGAGATTTTGGACCG	3654
Qy	3602	gatacagccacccagcagagagacgagccctccacgcgggcagcgctcaagttccacgtccagg	3661
Db	3715	CGATTGAGAAATCACACGATCTTTCAAGTAGTTAAGCATGAGC-----GAGAACTCGG	3768
Qy	3722	aaatgagagagcgcttcagtaacttgatataacttgaagtggagcagagagaaatagcgg	3721
Db	3769	CTCTGGTAGGTTACTGTGCAACCGCTCGAGGAGATCCAGGAGTTCCGATTAACCATTCG	3828
Qy	3782	ggccacagcagcgcgacagcagccgcgccttggacagcagcgtcacagtttgcgtgt	3841
Db	3829	CGCTACAGACCCAGGAAACGACCCAGGACCGCTGTCCACGACATGACGTTCCAGATTGT	3888
Qy	3842	cttcgtacccacacagggcgacccagttgttcagcgaanaatgacgtactgtcccttc	3901
Db	3889	TTTTGTGGCCACGCAAGAGACCTAGATTGCGGTCTCGAAMAATGCTGTCGTTTCAT	3948
Qy	3902	cgaaggtgaaagaagcgctccgtgagagttttgagcttgcgcgcaagcagaagacttaaaa	3961
Db	3949	AGAAAAAGATCCCGGATGGAAGAAGTCTACCAACTCTCTTAGCAACAATCAAMAA	4008
Qy	3962	ccacctctgcgaanaatgactgccaagaatctactcaagaatttatctggacgccaacaaga	4021
Db	4009	CCATCTCTGTGAAGACACACTGTCAACGATTTTCTGTATTTATTCGATGGCCAACAGCA	4068
Qy	4022	gggtcttctgtacttgcagcagatcaacgactcalctcccttgcgcgagagtgttgaacg	4081
Db	4069	AGGTCAATTTGGCCTGTGATCTCTGTGCAACAAGTGTTCCTGAGAAAGAGCTGATAG	4128
Qy	4082	cgaaggtggcagcgtcttaacagctgtgacatcgcgcgcgcgagcaactcccgagcgcactg	4141
Db	4129	GGAAACAAGTCCCTCCACACTCTGGAAGTGGGGGTGTAACTCTGCCGATG---GTGG	4185
Qy	4142	gataccctctgcagacttcaactcctctgttgaacggttcaatgtgaagaagaagaaacccg	4201
Db	4186	CATTTCACATTTCCTGCTTCCATCTTACTGTCACTGTTACCGTGAAGGAGGACACCTCG	4245
Qy	4202	cccaattctcgagcagacacttaccacaacgpcggcaatttcgaagtttgcacagacttggc	4261
Db	4246	TCCAGTGTGTTGTGAGGGAATTGTACACCGAAGGAAATCCACAGGGGACTCCATCGGAG	4305
Qy	4262	ggaactgtcttacttgcagggcgagacacacagaagaagacacacatcactacacataga	4321
Db	4306	AGAGCTCTCAGATTACATCTCGAACCCACAGTCTGAAGGCTCGCCATTACTTATGCTTAGA	4365
Qy	4322	ccggtgcgagatgcagcttgcagacagcgcttgaagcgcttgcgcacttgcgtctcgctc	4381
Db	4366	CTACGATACATGTTGTAATGAGACCCCGCTTGAGAGCGATGAGACAGTGGCTTTCTGACT	4425
Qy	4382	gcaatgcacacacggcgctgttcttgcctcaatltgcacccacccgcttccatgcagcgcat	4441
Db	4426	GAAAGCTCAAAACGGAGTGTGTAGCGTTATATTCACAGCCACAGCGCAAGATGATGACT	4485
Qy	4442	gttcagagttcgaagcgtacgtctacagatacagcttctgcaatcgcacacagccgttgtaa	4501
Db	4486	GTTCATAATTTGAAGTACACAGCTACTATACACGGCCGGCGCTGAGGACCGACCGACGTAC	4545

[illegible]

	APPLICATION NUMBER:	US/08/982,129	
	FILING DATE:		
	CLASSIFICATION:		
	PRIOR APPLICATION DATA:		
	APPLICATION NUMBER:	US 08/326,117	
	FILING DATE:	19-OCT-1994	
	ATTORNEY/AGENT INFORMATION:		
	NAME:	MILLMAN, ROBERT A	
	REGISTRATION NUMBER:	36,217	
	REFERENCE/DOCKET NUMBER:	7112-0037.00	
	TELECOMMUNICATION INFORMATION:		
	TELEPHONE:	(202) 887-1500	
	TELEFAX:	(202) 887-0763	
	TELEX:	90-4030	
	INFORMATION FOR SEQ ID NO: 1:		
	SEQUENCE CHARACTERISTICS:		
	LENGTH:	5577 base pairs	
	TYPE:	nucleic acid	
	STRANDEDNESS:	double	
	TOPOLOGY:	linear	
	MOLECULE TYPE:	cDNA	
	FEATURE:		
	NAME/KEY:	CDS	
	LOCATION:	197..4780	
	US-08-982-129-1		
Query Match	35.6%	Score 1954.6;	DB 3; Length 5577;
Best Local Similarity	63.9%;	Pred. No. 0;	
Matches 3248;	Conservative 0;	Mismatches 1769;	Indels 62; Gaps
Db	152 atcattgcaaaatgggggttagagagttcttccagcgagtactggtctcttaagctc	2111	
Db	187 ATATAGTGCATGGCAGTTGACGTGCCAATCGCTCCCTCGTGCTGTTATAGCGCC	246	
Oy	212 tgcgcgaactagaaccaacgatitlcttatatatagacaataaccaggcgaagactc	2717	
Db	247 TGCAGTTTTACTCTCAAGNAGATATGGGTATATACCGCCATCCCAAGGCTACCACAGCC	3066	
Oy	272 ggaac---tgccgcctatltga ltlacgaaggaanaalcatlgagltgaacagcctctaacc	328	
Db	307 GGATAATTTGGCAGTACTAAATTTTGAAGGCCAGACATGAGNGTCAGAGGCCCGCTGCC	3666	
Oy	329 cggccgcgaaccggagagaagtatgtatgagaact-----tcltacggatcaaat	3797	
Db	367 CGCCCCGGAGGGGATGACCTGTGCATGGAGCGCTTACCCAGTGTAACACCAACTCGG	426	
Oy	380 gatcaagtcatalacatlgaaggaagaanaatcgaaggaagcgtlcataltgcgaagcttaa	439	
Db	427 CACGCGAGTCTATCNATGNATGATGAAGAAGATGAAAGACAATCATCCCATACTTAA	4866	
Oy	440 ctatcaaggtgccaacacgcccgtgtcgtcgtatattatgtcaaggccagccccagaacct	4899	
Db	487 TTAT-AACGGACCATCAACTCCGTTCAATTGNAACHTYTTTATCCGGTTCTGTAATC	545	
Oy	500 ggagcccttgaagttgcagacgaatgaagcagaagcgcacaatlgagccctgttlaacgaag	5559	
Db	546 TGCATATCCCGGTATCATCAGAGAATTGACAACGGGAGTGCATC--TATATCATACGGAAG	6033	
Oy	560 acaaagactacagacagcaacacatgtcacagactatgtgttctcaatccaalgtagaaggtga	619	
Db	604 ACACGATTCACAGTAGTTGCCCGCGCAATGCAGCACATCATTTCAATGTGCGCGTGCAGCGCCA	6633	
Oy	620 atcagagccgtactgt	679	
Db	664 GTCCGTGTGTGCAGCGGTGTCTTCGCTATGTCAATCATATATACAAACGGGCCATCAT	7223	
Oy	680 gcaagttgtcagcgccctgtcgttaattccagaacatlgagcgaggtcgaactgaacgcgactcgt	739	
Db	724 ACAAAACTTCGAGCGCTTGCCGGGTTCTCGAAGCTGGGGAGCGCAAGGTGTAACAGATGCAC	78	
Oy	740 gtaccaaatgtcagaccgcgaacggtgaattcaatcagaccgcctcaatgaagttcgtgtcga	799	

OY	2951	cgaagtggacccctggacaactggtctgcctctgaagttgcattacccg-----ggagaca	3001
Db	2998	CTCCGTGAACCGAGAGACCGCGCTGGTGTACGTGGACTATGAGACCAAGGCTAGTGGCGA	3057
OY	3002	aaatctgacccgcgatggagatgagcctacatactactctgtgtaattccatcogtaa	3061
Db	3058	GGTGGTGGACCGGTGATGTGATTAACCAACGACGCGTACTCTTCAACGTCACTGACAA	3117
OY	3062	cttcctttctgatgtgtgacggttagagaaacaaagacgaagttgaaatatgtgcgtctc	3121
Db	3118	CTTCATGGGGGAGGAGGAAAGCTATACGAAGAAATCAGAACGACAGAAAGTTCTGTTATCTT	3177
OY	3122	attgagtgtgaacgaacaagcctccctagaatgcacatgtctgtatgaccccggtttgagt	3181
Db	3178	GTTGGATGTGAATGACAACTCTCTTAATTTGCCACCGCGGACGAACTCTCTTGACACT	3257
OY	3182	tcccgaaagagcaagttgtgtgtgtccgtgtactcccaagaatctacgcacgcgaaagga	3241
Db	3238	ATCTGAAACCTTAAAGCAGGCGCTCCGCTTGACCACTATCTTCCGCCGCGACCGCGA	3297
OY	3242	tgaaccgacaagaacgaacactcggttgcgttagaagaatccctgagccctcaagatcacga	3301
Db	3298	CGACCCCGACACACACACTCCAGGGCTGGCTACGAGATCTCGAACTC--AGCACGA	3354
OY	3302	ccgagacactgaggtgtgcggaatctctcaacatgaactcgaattgaaacaaacttggga	3361
Db	3355	GGCGGACATCGAAGTCCCGGAGCTGTTTGATGATACAGATCGGAAAGCTCACGGGAGA	3414
OY	3362	acttgagacgcgtctatgacttgaaggggtgatgtgggacattcgaatatctatgtagac	3421
Db	3415	GCTGGAGACCGCCATGGACCTCAAGGAGATATGGGGGAGTACGCTATACATATACGGGC	3474
OY	3422	cttcgacacagcgtcacccgcgagacagaggttccaaacgaagctacacccgtgtatccgcc	3481
Db	3475	ATTGACACCGGCACTTCCGAATATGTCCATGAACGAGAAATATGAGCTGATCATCTC	3534
OY	3482	ctacaacttccacacccctgtgtgtgtgtctccgcaaacccgaactcgtcatctggtcttc	3541
Db	3535	GTTCACTACTACACCGCCTGATGTCGTCTTCCCGACCAAGATGCGCTCATACGACTTGC	3594
OY	3542	taggagcgcgcgaacagaagcgcgtctctgacgagcgtgcgaacagatctcttgaagc	3601
Db	3595	GAGGGAACGACTGTATTCATGATGAGTGTACGGACAGTGAACGGAGAGTTCTTGGACG	3654
OY	3602	gactcaagccacgcgaaggaagggcctcgaagcgcggcagcgtcgaactcaccgtccaggt	3661
Db	3655	GATATCGCGACTGATATCCGACGGAGACTCCACGGGGCGTCTGCTACCTTCCAAGTGTAGG	3714
OY	3662	aaatgaggaagccgcttcaagcttctgataataactgtagagtgagacgaaggaataatagag	3721
Db	3715	CGATGAGGAATCACACGATCTTTCAAGTAGTTAAAGATGEC-----GAGAACCTGG	3768
OY	3722	gcgaactatattacgccaagcttttccagaagcaatcaagacaattcaagatcagacgcg	3781
Db	3769	CTCGTTGAGTTACTGCAACCGCTTCAGAGGAGATCAGGAGATTCCGATTAACATTCG	3828
OY	3782	ggcacaagcagcggcgcgcgaagcccgccgcgtcttgaccagcagtcacgttltgtgtgt	3841
Db	3829	CGCTACACACCGAGGAACGCCAGGACCGCTGTCCACGAGCATGACCTTGAGAGTTGT	3888
OY	3842	cttcgtaccacaagggcgagccacgagtgtctcagcgaanaatgacgtactcgtcctctc	3901
Db	3889	TTTTGTGCCCAACGAGAAGGAACTAGATTCCGGGTCTCTGAAACATGCTGTGCTTCAAT	3948
OY	3902	cgaaggtgaagaagcctcgtgtgagatlttgagctgcgcgaagcagaagaccttaaaa	3961
Db	3949	AGAAAAGATGCTCCGCGCATGTGAAGAGTCTACCACTTCCCTTAGCACACAGATCATMAAA	4008
OY	3962	coactctgcgaanaatgactgcgaagatctactcactcaagttatgtgaaggaacaga	4021
Db	4009	CCATCTCTGTGAACAACACTGTCTACACCAATTTACTGATTAATGATGGCAACGCA	4066
OY	4022	gggtcttcttgtaatgacagtaacgaagcagcgtcatctccctcttcgcagaggttggaccg	4081

[illegible]

```

RESULT 5
US-08-272-882D-1
; Sequence 1, Application US/08272882D
; Patent No. 5516685
; GENERAL INFORMATION:
; APPLICANT: Lichensteijn, Henri
; TITLE OF INVENTION: Isolation and Characterization of No. 5516685el
; TITLE OF INVENTION: Protease from Streptomyces Lividans
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Amgen Inc.
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320-1789
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/272.882D
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oleski, Nancy
; REFERENCE/DOCKET NUMBER: A-199A
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2261 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IS-08-272-882D-1

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2002, 16:00:42 ; Search time 61.22 Seconds
(Without alignments)
631.136 Million cell updates/sec

Title: US-09-715-909-2

Perfect score: 8909
Sequence: 1 MGVERFPAYLVLSASAL.....THGNFGRKSTPYLPQPHK 1717

Scoring table: BLOSUM62

Gapop 10.0 , Gapect 0.5

Seq: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents -AA: *
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep: *
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep: *
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep: *
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep: *
5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep: *
6: /cgn2_6/ptodata/2/1aa/Backfilest1.pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4693	52.7	1528	1	US-08-326-117B-2
2	4693	52.7	1528	3	US-08-982-129-2
3	459	5.2	119	1	US-08-326-117B-10
4	459	5.2	119	3	US-08-982-129-10
5	457.5	5.1	138	1	US-08-326-117B-12
6	457.5	5.1	138	3	US-08-982-129-12
7	398	4.5	146	1	US-08-326-117B-15
8	398	4.5	146	3	US-08-982-129-15
9	382.5	4.3	122	1	US-08-326-117B-14
10	382.5	4.3	122	3	US-08-982-129-14
11	376	4.2	120	1	US-08-326-117B-16
12	376	4.2	120	3	US-08-982-129-16
13	363.5	4.1	113	1	US-08-326-117B-8
14	363.5	4.1	113	3	US-08-982-129-8
15	352	4.0	124	1	US-08-326-117B-13
16	352	4.0	124	3	US-08-982-129-13
17	337.5	3.8	112	1	US-08-326-117B-17
18	337.5	3.8	112	3	US-08-982-129-17
19	324	3.6	1026	1	US-07-998-003A-95
20	324	3.6	1026	1	US-08-453-274B-95
21	324	3.6	1026	1	US-08-453-695A-95
22	324	3.6	1026	1	US-08-268-161A-95
23	324	3.6	1026	2	US-08-453-702A-95
24	324	3.6	1026	4	US-09-099-639-95
25	324	3.6	1026	5	PCT-US93-12588-95
26	324	3.6	1026	5	PCT-US95-08071-95
27	324	3.6	1203	1	US-07-998-003A-103

28	324	3.6	1203	1	US-08-453-274B-103	Sequence 103, App
29	324	3.6	1203	1	US-08-453-695A-103	Sequence 103, App
30	324	3.6	1203	1	US-08-268-161A-103	Sequence 103, App
31	324	3.6	1203	2	US-08-453-702A-103	Sequence 103, App
32	324	3.6	1203	4	US-09-099-639-103	Sequence 103, App
33	324	3.6	1203	5	PCT-US93-12588-103	Sequence 103, App
34	324	3.6	1203	5	PCT-US95-08071-103	Sequence 103, App
35	301.5	3.4	106	1	US-08-326-117B-9	Sequence 9, App1
36	301.5	3.4	106	3	US-08-982-129-9	Sequence 9, App1
37	298	3.3	104	1	US-08-326-117B-11	Sequence 11, App1
38	298	3.3	104	3	US-08-982-129-11	Sequence 11, App1
39	297.5	3.3	979	4	US-08-878-474-5	Sequence 5, App1
40	295.5	3.3	797	1	US-08-453-695A-112	Sequence 112, App
41	295.5	3.3	797	1	US-08-268-161A-112	Sequence 112, App
42	295.5	3.3	797	2	US-08-453-702A-112	Sequence 112, App
43	295.5	3.3	797	4	US-09-099-639-112	Sequence 112, App
44	295.5	3.3	797	5	PCT-US95-08071-112	Sequence 112, App
45	280.5	3.1	555	2	US-08-453-702A-98	Sequence 98, App1

ALIGNMENTS

RESULT 1
US-08-326-117B-2
Sequence 2, Application US/08326117B
Patent No. 5693491
GENERAL INFORMATION:
APPLICANT: BULLA, LEE A.
INVENTOR: JET TAB
TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
TITLE OF INVENTION: TOXIN
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESS: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/326,117B
FILING DATE: 19-OCT-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MILMAN, ROBERT A
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 7112-0037.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1528 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-326-117B-2

Query Match 52.7%; Score 4693; DB 1; Length 1528;
Best Local Similarity 60.7%; Pred. No. 0;
Matches 925; Conservative 220; Mismatches 357; Indels 22; Gaps 13;

Seq 1 MGVERFPAYLVLSASALANORCSYIAIPR-PPPELPPIDYEKSNSEOPPLIGPT 59
1 MAVDVIRNAFLVITIAVAVLAQKCGTWTATPRLPNDLFLVNFEGQTSQRLPAPE 60

QY 60 REEYOMENF---LPDOMIOVIMEEIEGDIYIAKLANTOGSNTPLVSI--MSGOPRAQLG 114
D 61 RDDLCMDAYHITANLGOVYIYMEDEIDEITIALINANGSTPIELPFLSGSNTL-LM 119
QY 115 PEFRONADGOMSLVITGRODYETATMOSYVESIOVEGSOAVLVALEIVINDNPILQ 174
D 120 PVIRR--VDNGSASHHNRHUYELPGMOQYMFNVAVDQSLVAGSLAIVINDNAPILQ 177
QY 175 VVSACVPEHEEARLDCVOYSDRDGELSTREMTFRVDSRADDESIFVWNGEYDPSDW 234
D 178 NFEORVPELDEPGLTECYQVSDADGRISTEFMTFRIDSVR-GDEEFYIETKRNIPQW 236
QY 235 FNMKTIVGINSPLNEETQOLHFSVTASDSLPNNHTVTMNOVENVESRPRMVFESVQ 294
D 237 MMLNMTIVGNTSLNFVISPPLHFSVTALDSLPNHTVTMNOVANNSRPRMVEIFAVQ 296
QY 295 QFDEKTNOSFLRAIDGDTGINRAINYTLIRDDADDFESLEVI--EDGAILHYTEIDRD 351
D 297 OFEERSYQNFVRAIDGDTINMPINRYLTINEEDTFESIEALPGKSGAVELVSPDRD 356
QY 352 KLERELMUTIVAAKSTPASFATEHIFIIYDNDVQDRPEPLHKEYSIDIMEETPMNTNF 411
D 357 TLQREVFLTVAAKYDEAFSTSTNVAVIIVTDINDQRPPIHKEYRLAIMEETPLTNF 416
QY 412 NEEFGRHDDGEMAOYVELEDFPPGASAFYIAPGSGYOROTFIMGITNHTMLDYED 471
D 417 DKERGFHKDGLQANAOYVRLSVDPGCAEAFYIAPVGVQOROTFIMGITNHTMLDYED 476
QY 472 VIFONIIRKAVADNNAHSHVGEALVYVNLINMDELPIFEBSYSASFKETVAGAPVA 531
D 477 PEFOSTIRVAVATNDNRHGVVALVHIDLINMDEQPIFEHVAQVTFDETEGGEFVA 536
QY 532 TVLADRIDDVVYHSLMGNAVDTLFDSESTGEFVSMDAFDYROMTLPVOVADPTL 591
D 537 KAVADRIDGVEHETLGLMNAVFLTLKLGDIKVSANDSNHRESELEVOVATPTL 596
QY 592 GDGPHNTVTOVLELEEDVNTPTLRLPRSTPSVEENVPEGEISREITADPTSAVL 651
D 597 GE-PHNTQSOLVIRLNDINNTPTLRLPRSGPOVEENVPOGHVITQGLRATDPTTADL 655
QY 652 WFEIDWSTWATKOGRETNPLEYVGCIVETIYPTEGNRGSAIGRLVVOEIRDNVTIDE 711
D 656 RFEIIMWDSFATKOGROANDEFRCVCIETIPEINNRRGLAIGHVAREIHNVTIDYE 715
QY 712 EPEMXYLVAVRDNTVYGGDYDEATFTIITIDMNDNAPFANGTLQTMVRELAASGT 771
D 716 EPEVSLTVAVRDNTVYGGDYDESMLTITITIDMNDNAPVWVEGTLEONFRERESAGL 775
QY 772 LIGSLATDIDGPLYNOVRYTIOPRNNTPEGLVKIDFTTGOIEVDANEAIDADEPFRYL 831
D 776 VVGSVRADIDGPLYNOVRYTIFPREDFDKOLIMDFLGOISVNTSADLADPTPRHL 835
QY 832 YTYVYASDECSLENTECPDPSNTEFEVGDIEIETIDNKNVPEPLTEKNTYVWENA 891
D 836 YTYVYASDRCTEPPADCPDPPTVWETETEGNTIHTIDNNKVPQAEKTKFTVYIYENA 895
QY 892 TSGDEVOLYSHDRDRDELHYVRYTMNFAVNPRLRDFEVDLDRGLEVHP--GDEK 948
D 896 THLDEVYLLIASDLDRDEIYHTVSIVYINAVNPRLMNEFSVNRKGLVYVDYETOGSSEV 955
QY 949 LDRDDEPTHTIEVNFIDFSDGDRNODEVELEFVLLVDVNDNAPMPPLDELRFVS 1008
D 956 LDRDDEPTHTIEVNFIDFSDGDRNODEVELEFVLLVDVNDNAPMPPLDELRFVS 1015
QY 1009 EGAVAGVAVLEIYAPDDEPTDNSRVYGLIDLTITDRDIEVDOLFMTMSIENKTEGL 1068
D 1016 ENLKGVARLEPHIFAPDDEPTDNSRVYGLIDLTITDRDIEVDOLFMTMSIENKTEGL 1074
QY 1069 ETAMDLRGYWGVEFFIYAFPHGVPOQSNMETYTVIRPNVHHVFPQDVSIRLSR 1128
D 1075 ETAMDLRGYWGVEFFIYAFPHGVPOQSNMETYTVIRPNVHHVFPQDVSIRLSR 1134

QY 1129 ERATEGVLATAANEFLPIYATDEDDGLHAGSVTFHVQNEEAVOYFDITEVGAGENSQ 1168
D 1135 ERAVINGVLAVNGEFLERISATIDPDGLHAGVTFVQVDEESQRYGV--VNDGENLS 1192
QY 1189 LILROLPEQIROFRITIRATDGEPEPLMTDVTFSVFVPTQDVPVSENAATVAFE 1248
D 1193 LRLQAVPEEIRREFRITIRADDOGTDPGLSTDMTFRVFPVPTQGEPPASSEHNAVIE 1252
QY 1249 GEEGLRESFEIPOAEDLNHLCEDDCODIYREFIDGNNEGFLVLDQSSNVLSLOEIDRE 1308
D 1253 KSAGEESHQPLADIDNHNLCEDDCCHSITRILIDGNSGEGHGLDVPYNNRFLPKKELRE 1312
QY 1309 VATSYTLHIAASNSPDATGPILOTSLIVTVYVREANRPJFEODLYTAGISTDSIGRE 1368
D 1313 QSASHTLQVAAANSBPG-GIPLPASILVTVYVREADRPVFRVRLYTAGISTDSIGRE 1371
QY 1369 LITVASHTEDDTITTTIDRASMOGLSSLEAVRDSAFALHATTGYLSLNMPTASMGAF 1428
D 1372 LRLHATQSESAITYYADYDTMVVDPSELVAVROSAPVLMQGTGLTINLPTATMGLF 1431
QY 1429 EEDVATATASADIDFARYKVVYLLISSQNRVTFIFDQLETVBOONRFIATSTGFNMTCN 1468
D 1432 KFEVATATAGAQDRTDVTYVYVSSQNNVYFVWTLQVEDNRDIFADITSAGFNMTCN 1491
QY 1489 IDQVVPFSDS--SGVAQDDTTEVRA 1511
D 1492 IDQVVPANDPVTGVALHSTQMAA 1515

RESULT 2
US-08-982-129-2
Sequence 2, Application US/08982129
Patent No. 6007981
GENERAL INFORMATION:
APPLICANT: BULLA, LEE A.
TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
TITLE OF INVENTION: TOXIN
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/982,129
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/326,117
FILING DATE: 19-OCT-1994
ATTORNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 7112-0037.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1528 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-982-129-2

Query Match 52.7%: Score 4693: DB 3: Length 1528;
 Best Local Similarity 60.7%: Pred. No. 0:
 Matches 925; Conservative 220; Mismatches 357; Indels 22; Gaps 13;

QY 1 MGVERFPVAVLVLASALANORCSYIIAIPR-PETPELPIDYEGKSMSEQLIPGPT 59
 1 MAVDVRLAFLVFLIAPVLAQRCGYMTAIPRLPRDNLPLVNFSGQWMSQRLPAPB 60
 60 REVCENENF---LPDMQVYIMEEIEGDVYIAKUNQGSNTPVUST--MSGQPAQIG 114
 61 RDCLCDAHVITANLGTQVYMEDEIEDEITAILNYPNSPTPELPFLSISYML-LM 119
 115 PERROMAOGMSLVITODDYETATMGOSVPSIOVSGSQAVALVLEIYNDNPPILQ 174
 120 PYLR--VNDGASHHAKHTELPQMOQMEVNRVDGSLAVAGSLAVINDNAPILQ 177
 175 VVSACVPEHGEARLDCVYOVSDDRGESITREMPRVDSRRADESIFVNGEYPSDM 234
 178 NPEPCVPELGEPLGECYQVSDADGRISTEMFRIDSVR-GDETFYIERTNIPNQM 236
 235 FNNKMTVGINSPLNETTQHLIFSVTASDLPNNHTVMQVENVESRPRKWEISVQ 294
 237 MMLNMTIGVNTSLNFTSPLHFSVTALDLPHTHTVMQVAVANSRPRMLEIFAVQ 296
 295 OPEKTNOSFSLAIDDDGINRAINYTLIRDDADDFSLVY---EDGAILHVTIDND 351
 297 OPEKSYQNFVYADIDGDEINMPYRLITMEEDFSEIALPGSKSAVFLVSPIDND 356
 352 KLERELNLTIVAKSSTDAFSAFETIAHIFLVNDVNDQREPLKHEKSIDIMEETPLNF 411
 357 TLQREVPILYAKKDEAFSTSTNVIIIVDINDQREPIHKEKRLAMETPLITNF 416
 412 NEBFGHDLGEBNAQYVLEEDVPRPGAASAFYIAPSGSGYORQFIMGTINHMLDYED 471
 417 DKEFGHDLGCONAQYVLESVDPGAFAFYIAPEVGYORQFIMGTINHMLDYED 476
 472 VIFONTIIVKAVDMNNAHGEALVYVNLIMNDELPIFEESYSASKEIVGAGFPA 531
 477 PEFQSTINIVATDNDNTHVGALVAILHIDLIMNDEQPIFEHAVQVITDETEGEEFFVA 536
 532 IVALADRIDDVVHSLMNAVDYLEIDESTGEIIFYSMDAFVIRHONTLFFQVRADTL 591
 537 KAVAHROIDGVVHRLGNAYNFIIDKLTGDIRKSANDSRYHRESLEFQVARTDPL 596
 592 GCGPHNTVTLVILEEDVNNPTPLRLRSPVSEENVPEGEYSIREITATDPDTSAVL 651
 597 GE-PFHATSQVIRLNDINNTPTPLRLRSPQVSEENVPDGHVITQELRATDPTADL 655
 652 WREIDMDSWATKQGEETPEVGCIVETIYPTBGNAGSAIGRLVQEIENDVITIDE 711
 656 REIIMDTSFATKQQAOPDEFRCVETIETFEINNGLAIGRVAREIHNHTVIDE 715
 712 EPEMLYLVTRVBDLVVDYDEATFTTIIIDMNDNAPFANGTLTOMRARELSAGT 771
 716 EEEVLSLTVRADNLVYDDYDESMLTTIIDMNDNAPVWEGTLEQVRYRESASAGL 775
 772 LIGSVLADIDGPLYQVRYTIQPRNTPPEGLVKIDFTTGQIEVDANEADADEPMREYL 831
 776 VVGSVYADIDGPLYQVRYTIFREDFTKDLIMIDFLGQISVMTSGAIDADTPRFL 835
 832 YXTVTLASDCSLENRECPDSNTEFEVDDIEIETIDNNKVPPEPLTEKFNFTVYWEHA 891
 836 YXTVVASDCSTEDPADCPDPTWETEGNTIHTIDNNKVPQAEFTKFDVYIYENA 895
 892 TSGDEVOLYSDRDDELVHTVRYTMNPAVNPRLDEFEVDLDGRLEVHP---GDEK 948
 896 THLDEVVTLIADLDRLDELVHTVRYTMNPAVNPRLNFSVAREGLVYVDETQSGSEV 955
 949 LDRDGEPTHTTFVNFIDNFSQDGRNQDEVEIFVLLVDVNDNAPEMPLRDELRFDS 1008
 956 LDRDGEPTHTTFVNFIDNFSQDGRNQDEVEIFVLLVDVNDNAPEMPLRDELRFDS 1015

QY 1009 EGAVAGRVLPETVAPDRDEPDNDNSRGVGLDILITRDIEVPLFTMISLENKGTGL 1068
 DB 1016 ENIKQGVRLPEPHFAPDRDEPDNDNSRGVGLDILITRDIEVPLFTMISLENKGTGL 1074
 QY 1069 ETAMDLRGVGTVEYLFIEAFDGGYPOQRSNETYTLVIRPNFHHFVFPPOPDSYRLSR 1128
 DB 1075 ETAMDLRGVGTVEYLFIEAFDGGYPOQRSNETYTLVIRPNFHHFVFPPOPDSYRLSR 1134
 QY 1129 ERATEGVATANEFLIETATDEDEGLHAGSVYTHVQSENEAVQYFDTTEVGAGENSQ 1188
 DB 1135 ERVINGVATANEFLIETATDEDEGLHAGSVYTHVQSENEAVQYFDTTEVGAGENSQ 1192
 QY 1189 LILRQLEPQIOKROFRTITATGTEPEGLMDVDFSVYVFPQGGPVSNAATVAFPE 1248
 DB 1193 LRLQAVPEIEIEFRITATDGGIDPGPLSDMTFRVYVFPQGGPVSNAATVAFPE 1252
 QY 1249 GEGRLSEFELPQAEDEKLNHLCEDDQDIYRFRIDGNBGLVLDQSSNVISLAQELDRE 1308
 DB 1253 KSGAMESSHQPLAQDKNHLCEDDCHSYRRIIDGNSGHRGLDPRVNRRLFKKELLRE 1312
 QY 1309 VATSTYLHIAASNPDAIGIPIQTSILVTVVNRANRPRITEQDLTYAGISTLSIGRE 1368
 DB 1313 QGASHTLQVAASNSPDG-GIPLASITLVTVVNRANRPRITEQDLTYAGISTLSIGRE 1371
 QY 1369 LILVRASHTEDDTITVITIDRASMOGLDSLEAVRDSAFALHATGVLSIMQPTASHMGH 1428
 DB 1372 LRLHAQTOSESAITVYADIDTVNVDPDSLEAVRDSAFALHATGVLSIMQPTASHMGH 1431
 QY 1429 EFDVATIDFASADIDTARVAVYLISSONRYTFLENDQLETVENRNFIAFTSTGNMTCN 1488
 DB 1432 KEVYATIDFAGADRDVYVYVSSONRYTFLENDQLETVENRNFIAFTSTGNMTCN 1491
 QY 1489 IDQVVPFSDS-SGVAODTTEYVA 1511
 DB 1492 IDQVVPANDPVYVVALEHSTQMAA 1515

RESULT 3
 US-08-326-117B-10
 ; Sequence 10, Application US/08326117B
 ; Patent No. 5693491
 ; GENERAL INFORMATION:
 ; APPLICANT: BULLA, LEE A.
 ; TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
 ; NUMBER OF SEQUENCES: 26
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MORRISON & FOERSTER
 ; STREET: 2000 Pennsylvania Ave. N.W.
 ; CITY: Washington
 ; STATE: DC
 ; COUNTRY: USA
 ; ZIP: 20006-1812
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/326,117B
 ; FILING DATE: 19-OCT-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MILLMAN, ROBERT A
 ; REGISTRATION NUMBER: 36,217
 ; REFERENCE/DOCKET NUMBER: 7112-0037.00
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 887-1500
 ; TELEFAX: (202) 887-0763
 ; TELEX: 90-4030
 ; INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-326-117B-10

Query Match 5.2%; Score 459; DB 1; Length 119;
Best Local Similarity 71.4%; Pred. No. 1e-31;
Matches 85; Conservative 15; Mismatches 19; Indels 0; Gaps 0;

QY 393 LAKEYSIDIMETPTLNFNEGFHDLGEMAOYVLELDPFGAASAFYIAPGSGY 452

DB 1 IHKEYRLAIMEETPLTNFDEKFGHDKLQGMAYVRLSEVDPGAALFYIAPGVGY 60

QY 453 ORQFTMGTHNTHMDYEDVIFQNIIRKAVDMNASHGALVYVLIIMNDLPIF 511
DB 61 ORQFTMGTHNTHMDYEDVIFQNIIRKAVDMNASHGALVYVLIIMNDLPIF 119

US-08-326-117B-10
Sequence 10, Application US/08982129
Patent No. 6007981

GENERAL INFORMATION:
APPLICANT: BULLA, LEE A.

TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
TITLE OF INVENTION: TOXIN
NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESSES:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W.
CITY: Washington
STATE: DC

COUNTRY: USA
ZIP: 20006-1812

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/982,129
FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/326,117
FILING DATE: 19-OCT-1994
TORNNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A.

REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 7112-0037.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763

TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-982-129-10

Query Match 5.2%; Score 459; DB 3; Length 119;
Best Local Similarity 71.4%; Pred. No. 1e-31;
Matches 85; Conservative 15; Mismatches 19; Indels 0; Gaps 0;

QY 393 LAKEYSIDIMETPTLNFNEGFHDLGEMAOYVLELDPFGAASAFYIAPGSGY 452
DB 1 IHKEYRLAIMEETPLTNFDEKFGHDKLQGMAYVRLSEVDPGAALFYIAPGVGY 60

DB 1 IHKEYRLAIMEETPLTNFDEKFGHDKLQGMAYVRLSEVDPGAALFYIAPGVGY 60
QY 453 ORQFTMGTHNTHMDYEDVIFQNIIRKAVDMNASHGALVYVLIIMNDLPIF 511
DB 61 ORQFTMGTHNTHMDYEDVIFQNIIRKAVDMNASHGALVYVLIIMNDLPIF 119

US-08-326-117B-12
Sequence 12, Application US/08326117B
Patent No. 5693491

GENERAL INFORMATION:
APPLICANT: BULLA, LEE A.

TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
TITLE OF INVENTION: TOXIN
NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESSES:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W.
CITY: Washington
STATE: DC

COUNTRY: USA
ZIP: 20006-1812

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/326,117B
FILING DATE: 19-OCT-1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A.
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 7112-0037.00

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030

INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 138 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-326-117B-12

Query Match 5.1%; Score 457.5; DB 1; Length 138;
Best Local Similarity 64.7%; Pred. No. 1.8e-31;
Matches 90; Conservative 22; Mismatches 22; Indels 5; Gaps 2;

QY 618 RLPSTPSVENVPEGEISREITATDPDTSAYLMFELIDWSTWATKQGRNTPEYVC 677

DB 1 RLPSTPSVENVPEGEISREITATDPDTSAYLMFELIDWSTWATKQGRNTPEYVC 60

QY 678 IYIETIYPTL-----GNRSATGRLVVOEIRDNVTIDFEFEMLYLVVRDINTVIGDY 733

DB 61 IYIETIYPTL-----GNRSATGRLVVOEIRDNVTIDFEFEMLYLVVRDINTVIGDY 119

QY 734 DEATFTITIDMNDNAPIF 752

DB 120 DESMLTITIDMNDNAPY 138

US-08-982-129-12
Sequence 12, Application US/08982129
Patent No. 6007981
GENERAL INFORMATION:
APPLICANT: BULLA, LEE A.

ATTORNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 7112-0037.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 146 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-982-129-15

Query Match 4.5%; Score 398; DB 3; Length 146;
Best Local Similarity 53.4%; Pred. No. 2,4e-26;
Matches 78; Conservative 14; Mismatches 26; Indels 28; Gaps 1.

998 P L P P E L R D V S E G A V A G R V L P E L I A P R D E P D T N S R V G Y G I L D L T I ----- 1045
1 P P P S L W T I S E N I K G G R L E P H I F A P R D E P D T N S R V G Y E I L L S T E R D I E V P E L F Y M 60
D B

OY 1046 -----T R D I E V P L F T M I S E N K T G E L T A M D L R G V G T Y E I F E A R D 1089
D B 61 I Q I I A N V T G Y E I L N S T E R D I E V P E L F Y M I Q A N V T G E L T A M D L K G Y G T A I Y I L A D 120
OY 1090 H G Y P Q R S N E T Y T V I R P Y N F H P V F 1115
D B 121 H G I P Q S M N E T Y E L I H P N Y A P E F 146

RESULT 9
US-08-326-117B-14
Sequence 14 Application US/08326117B
Patent No. 5693491
GENERAL INFORMATION:
APPLICANT: BULLA, LEE A.
APPLICANT: JI, TAE
TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
TITLE OF INVENTION: TOXIN
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/326,117B
FILING DATE: 19-OCT-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 7112-0037.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0763
TELEX: 90-4030
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 122 amino acids
TYPE: amino acid
STRANDEDNESS: single

[illegible]

Db 2 TTKFDTVYIYNATHLDEVTLLASDLDRBEIYHMSVYIYAVNPRMLNPFVSRENG 61
QY 938 RLEVIYIPGD-EKLDGDGDEPHITIVNFIDNFFSDGGRNODVEIIVLLDVNDAP 996
Db 62 LYYDYETGSGLDGDEPHITIRIFENLIDNMGEGENRQNDREYVILLDVNDAP 121
QY 997 M 997
Db 122 L 122

RESULT 11

US-08-326-117B-16
; Sequence 16 Application US/08326117B
; Patent No. 5693491
; GENERAL INFORMATION:
; APPLICANT: BULLA, LEE A.
; APPLICANT: JI, TAE
; TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
; TITLE OF INVENTION: TOXIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/326,117B
; FILING DATE: 19-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MILLMAN, ROBERT A
; REGISTRATION NUMBER: 36,217
; REFERENCE/DOCKET NUMBER: 7112-0037.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-326-117B-16

Query Match 4.2%; Score 376; DB 1; Length 120;
Best Local Similarity 65.6%; Pred. No. 1.3e-24;

Matches 80; Conservative 12; Mismatches 28; Indels 2; Gaps 2;

QY 1116 VFPPDSVIRLSRRATEGGVLAATANELEPIYATDEDDGLHAGSVTFHVGNEAVOYF 1175
Db 1 VFPTNDAVIRLARERAVINGVLAATVNGEFLERISATDPDGLHAGVTFQY-GDESGRYF 59
QY 1176 DITEVAGENGSQLILNQLPEQIROPFRITIRATDGGTEPGPLMTDVTFSVFPPTGDP 1235
Db 60 QVVD-NNGENLGSRLQLQAVPEELREFRITIRATDQGTDPGLSTDMTFRVFVPTGEP 118
QY 1236 VF 1237
Db 119 RF 120

RESULT 12
US-08-982-129-16

; Sequence 16 Application US/08982129
; Patent No. 6007981
; GENERAL INFORMATION:
; APPLICANT: BULLA, LEE A.
; APPLICANT: JI, TAE
; TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
; TITLE OF INVENTION: TOXIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Ave. N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/982,129
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/326,117
; FILING DATE: 19-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: MILLMAN, ROBERT A
; REGISTRATION NUMBER: 36,217
; REFERENCE/DOCKET NUMBER: 7112-0037.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; TELEX: 90-4030
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 120 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-982-129-16

Query Match 4.2%; Score 376; DB 3; Length 120;
Best Local Similarity 65.6%; Pred. No. 1.3e-24;
Matches 80; Conservative 12; Mismatches 28; Indels 2; Gaps 2;

QY 1116 VFPPDSVIRLSRRATEGGVLAATANELEPIYATDEDDGLHAGSVTFHVGNEAVOYF 1175
Db 1 VFPTNDAVIRLARERAVINGVLAATVNGEFLERISATDPDGLHAGVTFQY-GDESGRYF 59
QY 1176 DITEVAGENGSQLILNQLPEQIROPFRITIRATDGGTEPGPLMTDVTFSVFPPTGDP 1235
Db 60 QVVD-NNGENLGSRLQLQAVPEELREFRITIRATDQGTDPGLSTDMTFRVFVPTGEP 118
QY 1236 VF 1237
Db 119 RF 120

RESULT 13

US-08-326-117B-8
; Sequence 8 Application US/08326117B
; Patent No. 5693491
; GENERAL INFORMATION:
; APPLICANT: BULLA, LEE A.
; APPLICANT: JI, TAE
; TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS
; TITLE OF INVENTION: TOXIN
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER

Query Match 4.0%; Score 352; DB 1; Length 124;
Best Local Similarity 56.2%; Pred. No. 1.6e-22;
Matches 68; Conservative 17; Mismatches 36; Indels 0; Gaps 0;
QY 755 GYLTQTRVRELAASGTLIGSVLATDIDGPLYNQVRYTIQPRNNPEGLVKIDFTTGQIE 814
||| | |||:| | :||| | ||| ||| | | : | : | : |
Db 3 GTLEQNFRVREMSAGGLVGSVRADDIDGPLYNQVRYTIFPREDDTKDLIMELPHGSNF 62
QY 815 VDANEALDADEPWRFYLYTYVIADECSELENTECPDSDNYFEVPGDIELEIIDTNNKYP 874
||| | |||:| | :||| | ||| ||| | | : | : | : |
Db 63 REHKRIDANTPPRHLYTYVVASDRGSTEDPADCPPDPYMETEGNITITHIDTNNKYP 122
QY 875 E 875
|
Db 123 Q 123

Search completed: January 15, 2002, 17:20:45
Job Name: 4803 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2002, 17:12:57 ; Search time 108.48 Seconds
(without alignments)
2315.170 Million cell updates/sec

Title: US-09-715-909-2

Perfect score: 8909

Sequence: 1 MGVERFPFPAVLVSLASAL.....THGNNGFKSPYLPQHPK 1717

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

hed: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP:REMBL_17.*
2: SP:archaea.*
3: SP:bacteria.*
4: SP:fungi.*
5: SP:human.*
6: SP:invertebrate.*
7: SP:mammal.*
8: SP:mhc.*
9: SP:organelle.*
10: SP:plant.*
11: SP:rodent.*
12: SP:virus.*
13: SP:vertebrate.*
14: SP:unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	5607.5	62.9	1715	09xy09	09xy09 bombyx mori
2	5597.5	62.8	1716	09ndm5	09ndm5 bombyx mori
3	5595.5	62.8	1716	09ndm4	09ndm4 bombyx mori
4	5594.5	62.8	1716	09ndm3	09ndm3 bombyx mori
5	5429.5	55.8	1717	09gpb9	09gpb9 manduca sex
6	4631	52.0	1728	09h137	09h137 manduca sex
7	1039.5	11.7	2005	09vfh5	09vfh5 drosophila
8	818.5	9.2	3322	099nh1	099nh1 mus musculus
9	818.5	9.2	3354	099pe4	099pe4 mus musculus
10	758.5	8.5	2552	09h4k9	09h4k9 homo sapien
11	758.5	8.5	3354	09h251	09h251 homo sapien
12	622	7.0	1374	09vgs5	09vgs5 drosophila
13	616.5	6.9	3503	024292	024292 drosophila
14	595	6.7	1820	09vvg0	09vvg0 drosophila
15	593	6.7	4587	09qxa3	09qxa3 mus musculus
16	585	6.6	4349	09vnt08	09vnt08 homo sapien
17	579.5	6.5	4643	09vwt1	09vwt1 drosophila
18	574	6.4	4589	09w10	09w10 rattus norv
19	560	6.3	4351	088277	088277 rattus norv

20	530	5.9	3579	5	09vsn8	09vsn8 drosophila
21	519	5.8	2923	4	09hcu4	09hcu4 homo sapien
22	503	5.6	2920	11	09r0m0	09r0m0 mus musculus
23	498.5	5.6	2809	5	061230	061230 lytechinus
24	494.5	5.6	3312	4	09nq07	09nq07 homo sapien
25	492	5.5	1311	5	09tyw6	09tyw6 caenorhabd
26	484.5	5.4	1439	5	09xmw6	09xmw6 caenorhabd
27	483.5	5.4	3313	11	088278	088278 rattus norv
28	452	5.1	3034	11	035161	035161 mus musculus
29	448.5	5.0	1943	11	099pj1	099pj1 mus musculus
30	444.5	5.0	3014	4	09nq06	09nq06 homo sapien
31	442.5	5.0	1310	4	09bye9	09bye9 homo sapien
32	442	5.0	1181	4	09v526	09v526 homo sapien
33	434.5	4.9	4307	5	019319	019319 caenorhabd
34	405	4.5	2163	5	001912	001912 caenorhabd
35	403	4.5	2610	5	019482	019482 caenorhabd
36	374.5	4.2	3337	5	09rtw4	09rtw4 caenorhabd
37	369.5	4.1	1965	2	09v504	09v504 synecocyst
38	366.5	4.1	2204	5	09veu1	09veu1 drosophila
39	353	4.0	2240	5	09gr86	09gr86 drosophila
40	352.5	4.0	1340	4	09bz28	09bz28 homo sapien
41	352	4.0	1563	5	09vaf5	09vaf5 drosophila
42	344.5	3.9	829	4	09y5d4	09y5d4 homo sapien
43	344.5	3.9	932	4	09y5h0	09y5h0 homo sapien
44	344	3.9	1203	4	09hc56	09hc56 homo sapien
45	341.5	3.8	1337	4	09bz26	09bz26 homo sapien

ALIGNMENTS

RESULT 1
ID 09xy09 PRELIMINARY: PRT: 1715 AA.
AC 09xy09;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE BTR175 PRECURSOR.
GN BTR175.
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditylsia;
OC Bombycoidea; Bombycidae; Bombyx.
OX NCBI_Taxid=7091;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MIDGUT;
RA Nagamatsu Y.;
RT "Cloning, Sequencing, and Expression of the Bombyx mori Receptor for
RT Bacillus thuringiensis Insecticidal CryI(a) Toxin.";
RT Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=MIDGUT;
RA MEDLINE=98276887; PubMed=9614703;
RA Nagamatsu Y., Toda S., Koike T., Miyoshi Y., Shigematsu S., Kogure M.;
RT "Cloning, sequencing, and expression of the Bombyx mori receptor for
RT Bacillus thuringiensis insecticidal CryI(a) toxin.";
RT Biosci. Biotechnol. Biochem. 62:727-734(1998).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: TO THE CADHERIN FAMILY.
DR EMBL, AB026260; BAA77212.1; .
DR HSP, P15116; INCI.
DR InterPro: IPR002126; Cadherin.
DR Pfam: PF00028; cadherin; 6.
DR SMART: SM00112; CA; 10.
DR PROSITE: PS00232; CADHERIN_1; 1.
DR PROSITE: PS50268; CADHERIN_2; 10.
KW Calcium-binding; Cell adhesion; Glycoprotein; Repeat; Signal;
KW Transmembrane.
FT SIGNAL 1 21 POTENTIAL.
SQ SEQUENCE 1715 AA; 19335 MW; 2EC68401FB054C0F CRC64;

Query Match 62.9% Score 5607.5; DB 5; Length 1715;
 Best Local Similarity 63.5%; Pred. No. 0; Mismatches 377; Indels 29; Gaps 12;
 Matches 1095; Conservative 224; Mismatch 377; Indels 29; Gaps 12;

1 MGVEFPFVAVLVSLASLANQRCYIIAIPRPETPELPIDYEGKSMSEOPILGPTR 60
 1 MGVDRIILATLILIAEYVLAQERGFMAIPRPRLPELDFSGQTMQSLPLPADR 60
 61 EYVCMENF---LPDMQIYIWEEBIEGDYIIAKIANYGSGNTPLV--SMSGOPRAQLCP 115
 61 EDVCMGDMHAMPTGTGTQIYIWEEBIEGEVPLAKLNTGRPNVYIEPAPLISGSFNLVAV 120
 116 EPRQWAGQWGLVTRQDQYETATMOSYVFSIOYEGSQAALVLALEYINIDNPILQV 175
 121 IRRIPDSNGEMLHITQKODYETPGMOQYVFNIRIDGETLVAGVSLIYINIDNAPIIDA 180
 176 VSACVYIPEHGAARLDCVQVYSDRDEISTREPMTRFVSSRAADESIFYMGEYDPSDMF 235
 181 LEPQVDELGEARLFEQYVVTADAGRISTQFMQFRIDSR-GDCKIFYIGCANIPGEMI 239
 236 NKKMTVGINSPLNETTQIHLFSVTSASDSLNNHTVTMMQVYENVSRRPRAVEIFSVQ 295
 240 RMTMTVGINEPLNETNPLHIFSVTALDSLPTHTVTLVQVYENVEHRRPRAVEIFAVQ 299
 296 PEKTNOSPLRAIDGDRGINRAINVTILRRDADDFEFLVTE--DGAIIHVTEDRDX 352
 300 PEKTAQSPVPAIDGDRGINRAINVTILRRDADDFEFLVTE--DGAIIHVTEDRDX 359
 353 LERELFNLTIVAKSTDSAFETAHFIIVNDVNDQRPRLKHEYSIDIMEPTMLNPN 412
 360 LQREVFOGLIAIKYKDNESATANVILVNDQRPRLKHEYSIDIMEPTMLNPN 419
 413 EBFQGHDRDLGNAQYVLEEDYFPPGASAFYIAPGSGYQKOTFMGINTMTLDYEV 472
 420 QBFQGHDRDLGNAQYVLEEDYFPPGASAFYIAPGSGYQKOTFMGINTMTLDYEV 479
 473 IFQNTIIRKAVDMNNAHVGALVYVNLINNDDELPIEESYSASFKEYGAGPVPAT 532
 480 EFQRIIRKAVDMNNAHVGALVYVNLINNDDELPIEESYSASFKEYGAGPVPAT 539
 533 VLALDRIDIDVYVSLMGNAVDFLEIDESTGEIFVSMDDAFDYHRONTLFEVORADDTLG 592
 540 VRAHRDIDIDVYVSLMGNAVDFLEIDESTGEIFVSMDDAFDYHRONTLFEVORADDTLG 599
 593 DGRHNTVTLQVLEEDVNTPTLRLPRSPVSENVPEGEISELITATPDPTSAVLM 652
 600 E-PEHPATSQLIHLIEDVNTPTLRLPRSPVSENVPEGEISELITATPDPTSAVLM 658
 653 FEIDMSWTATKQGRETNPTEYGVCIETIYPTTEGNRGSAGRLVAQVEIRDNVTIDFEE 712
 659 FEIDMTTSATQGRANIEFNHNCVEIETIYPALNNRGSAGRLVAQVEIRDNVTIDFEE 718
 713 FEMULYVNVKRLDNLNVIGDYDEAFETIITIDMDNAPIFANGTILQTKRVRELAASGLT 772
 719 FEMULYVNVKRLDNLNVIGDYDEAFETIITIDMDNAPIFANGTILQTKRVRELAASGLT 778
 773 TGSVATLIDIDGLYVNVKRTIQPRNTPREGLVKIDFTTQOIEVDANEAIDADEPFAFYLY 832
 779 TGSVATLIDIDGLYVNVKRTIQPRNTPREGLVKIDFTTQOIEVDANEAIDADEPFAFYLY 838
 833 YVIVASDECSLNNRTECPDSNVFEVPGDIEIETIDNTNKNVPEPLTEKNTYVYVENAT 892
 839 YVIVATDRCTAEDPPDCPPDPTTWTETPGOVVIOIITNNKIFQPEIDOKRAVYITIEDAY 898
 893 SGDEVVQVLSHRRDDELHYVRYTMMFANVPRDLFEVDLDTGRLLEVYPGDEKLDR 952
 899 SGDEVVQVLSHRRDDELHYVRYTMMFANVPRDLFEVDLDTGRLLEVYPGDEKLDR 958
 953 GDEPHTITVNFIDNFSQDGRANDVEITVLLDVANDNAPEMALPDELPRDVSAGAV 1012
 959 GDEPHTITVNFIDNFSQDGRANDVEITVLLDVANDNAPEMALPDELPRDVSAGAV 1018

1013 AGVRLPEIYAPDRDEPPTDNRVGYGLIDLTITDRDIEVPLFTMISLENKTEGETAM 1072
 1019 KTRIQPHIYAPDRDEPPTDNRVGYGLIDLTITDRDIEVPLFTMISLENKTEGETAM 1078
 1073 DIRGYWEIYEIYAPDRDHCYPOORSNETYTVIYRPNHHVYFVRPQDVSIRLSREAR 1132
 1079 DIRGYWEIYEIYAPDRDHCYPOORSNETYTVIYRPNHHVYFVRPQDVSIRLSREAR 1138
 1133 EGVGLATANEFLEPLVATDEDEGLAGSVTFHVGNEEVAOVYEDTTEVAGENSQOLIR 1192
 1139 VNGLATVDEGEFLNRIYATDEDEGLAGSVTFHVGNEEVAOVYEDTTEVAGENSQOLIR 1196
 1193 QLEPEIQRFRTIRATIGTEPGLMTDVTFSSVYFVPTQDVSSENAATVAFEGEGC 1252
 1197 QLEPEIQRFRTIRATIGTEPGLMTDVTFSSVYFVPTQDVSSENAATVAFEGEGC 1256
 1253 LRESELPQAEIDLKHLCEDCODIYRFRIDGNNEGLFVLDQSSNVISLAQELDEVATS 1312
 1257 LRESELPQAEIDLKHLCEDCODIYRFRIDGNNEGLFVLDQSSNVISLAQELDEVATS 1316
 1313 YTLHTAASNPATGPIQTSILVTVNVRANPRPIEBODLYTAGISPLDSIGRELTLY 1372
 1317 YTLHTAASNPATGPIQTSILVTVNVRANPRPIEBODLYTAGISPLDSIGRELTLY 1374
 1373 RASHEEDTITVTRIDRASQOLDSLEAVRDSAFALHATGYVLSLNMQPTASMHGMEFDY 1432
 1375 QATHSENAPITITDMSVMTVPTLASVRETAFLINPHGVTLINQPTASMHGMEFDY 1434
 1433 IATDASAIDTARVKKYLLISSONRVTFIFDNOLETVNORNFIAATFSGFNMTCIDOV 1492
 1435 VATDPAGYSDRANVKKYLLISSONRVTFIFDNOLETVNORNFIAATFSGFNMTCIDOV 1494
 1493 VPESSSGVADDTVEVRAHPIRDVNPVQAEVAVRSQTVLLRTQLMLSTNSVLQDL 1552
 1495 VPESSSGVADDTVEVRAHPIRDVNPVQAEVAVRSQTVLLRTQLMLSTNSVLQDL 1554
 1553 VTGDPPTLEESQOIAVLAALSAVGLFCLVLALEPCRTALNRQALSMTRYGSV 1612
 1555 VTGDPPTLEESQOIAVLAALSAVGLFCLVLALEPCRTALNRQALSMTRYGSV 1613
 1613 DGLNRAGL-AEGTNKHAVGSGNPMKNEAIKAPDPDAIDASGDSPLTIDEMPOFRDY 1671
 1614 DGLNRAGL-AEGTNKHAVGSGNPMKNEAIKAPDPDAIDASGDSPLTIDEMPOFRDY 1673
 1672 FPPGDTSSSGIYVLMHGEATDNKP---VTHGNGFGKSTPLYPO 1713
 1674 FPPGDTSSSGIYVLMHGEATDNKP---VTHGNGFGKSTPLYPO 1708

RESULT 2
 09NDN5 PRELIMINARY; PRT: 1716 AA.
 AC 09NDN5:
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, last annotation update)
 DE CADHERIN-LIKE MEMBRANE PROTEIN (FRAGMENT).
 GN BTR175A.
 OS Bombyx mori (Silk moth).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Bombycoidea; Bombycidae; Bombyx.
 OX NCBI_TaxID=7091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=MIDGUT;
 RA Ikawa S., Tsuda Y., Fukada T., Sugimoto K., Himeeno M.,
 RT "cDNA cloning of the CryIa receptor variants from Bombyx mori and
 RL their expression in mammalian cells.";
 CC Submitted (Apr-2000) to the EMBL/Genbank/DBD databases.
 DR EMBL; AB041508; BAA99404.1; -.

DR InterPro: IPR0002126; Cadherin.
DR Pfam: PF000028; cadherin; 6.
DR SMART: SM00112; CA; 10.
DR PROSITE: PS00232; CACHERIN_1; 1.
DR PROSITE: PS50268; CACHERIN_2; 10.
KW Calcium-binding; Cell adhesion; Glycoprotein.
FT NON_TER 1716 1716
SQ SEQUENCE 1716 AA; 193412 MW; D704BFEFD3B39ADD CRC64

Query Match	62.8%	Score 5597.5	DB 5	Length 1716
Best Local Similarity	63.4%	Pred. NO. 0		
Matches 1094	Conservative 224	Mismatches 378	Indels 29	Gaps 12

QY	1	MGVREPPAVLVYSLASAALANCSVTIAPREPELPEIDVEXGWSMOPILPEPT	60
Db	1	MGVDRILATLTLILYLAETVLAAQERCGRMVAIPRRPRDLPELDEEGOTWSORLIPADR	60
C	61	EEVCMENF--LPDOMIOVIYMEEEIEGDVIAAKLANTOGSNPVL--SMSGOPRAOLGP	115
QY	116	EEFRNENADGOMSTVITQROXYETATMOGYSPTIOVEESQAVLALETVINDDNPPILQV	175
Db	61	EDVCMGCVHMTPTVYQTOIYMEEEIEGEVPIAKLNRGPVPIEPAFGLSPNLLVPV	120
Db	121	IRRLPDSGHEHLITIRODETPEGMQOYFNINIDETVLVAGVSLITVINDDAPLITQA	180
QY	176	VSACVPIBEGHARLTDQVVOYSDRGEISTRFMRFRYDSSRADESIFYMWGEPSDFM	235
Db	181	LEPCQVDELGBARLTCECVVYVTDADGRISIQPMQFRIDSR-GDDKLFYIQGANIPBEM	239
QY	236	NMKTAVGINSPLNETTQOLHIFSTVASTDSLPHNHTVTMAQVENESRPPRWAEISVQO	295
Db	240	RMUTVAVINEPLNETPEPLHIFSTALDTSLPHTHTVLAMOVENVERHPRWAEIFAVQO	299
QY	296	FDEKTNOSFSRALDGGTGINRAINYLLRIDDADDFSLSEIYE--DGAHLHTEIDRDK	352
Db	300	FDEKTNOSFPRAIDGGTGINRPHYLELAEEDTFPHITTEGGRGALLIYDPRIDRT	359
QY	353	LEBELFMILTAAKSTDAPEAHIFIIYNDVNDQRPDLHKEYSIDIEETPMTLNFN	412
Db	360	LOREVPOLSTIAKYDNESSATANVYIYNDINDQRPDLKEYRLINIMEETALTNFD	419
QY	413	EEEFHNRDIOGENAQYVELLEDVPRPAAAFYIAPBSGQROTFMGTLNHNMLDEYV	472
Db	420	QEEFEBHRDIOGNAQYVRELESDPADAAAFYIAPBVGQROTFMGTLNHNMLDEYV	479
QY	473	IFQWIIKIVAVYADNNNAHSHGEALVYNLLINMDELPIFEESGSASFKETVAGFVPAT	532
Db	480	EFGQIRLRVIATDMDNEHGVAVYVILNLNMWDEBPIFESHQVONSFKTEBKGFPVAN	539
QY	533	VIALDRIDIDVYVHSLMGNAVDLFIDESIGELFVSKMDAFDYNHONTLFFVOYRADDTLG	592
Db	540	VRAHDRIIDRIVEHTLGMNANNYLSIDKDGDIHVODDFDVHROSELFFVOYRADDTLG	599
QY	593	DGPHNYTQVTOLELEDVNNPTPLRLRPSSTPYEENVPBGYEISREITFADPTDTSAYLM	652
Db	600	E-PRHTATSOQLIHLEDINTPPLRLRPSRSPVNEEVPBGYIITSIRATDPTDTHAELR	658
QY	653	FEIDMDSTMAKQGRETNPEYVGCIVETIYFEGNGSAGIAGLVVQELRDVYIDFEE	712
QY	713	FEMLYLVARBDINTVYIGDYUDATFPIITIDNNNDAPFRANGTLOTMAYRELASGTL	772
Db	719	FEMLYLVARBDINTVYIGDYIDSTFPIITIDNNNDNPLWPGTLEOSLKVRRMSDAGVY	778
QY	773	IGSVLATIDGPLYNOVRYTIQPRNNTPEGLVYIDFTGQIEVDANEADADPEMRFYLT	832
Db	779	IGTLTADIDIGPLYNQRYRMKANEGRPENLIMIDFYTGITKTSGALIDADVPRRYNLX	838
QY	833	XYVIASBESLKNRECEPPSPSNFEVYBGDIEIELIDTNKAVPEPLIEKFTIYYVWENAT	892
Db	839	XYVAVADRASYAEBDDCPDPPTWETPPGVVJIDITNNKIPPEPDIQFAVYVYIEDAV	898

[illegible]

SEQUENCE FROM N.A.
 RP TISSUE=MIDGUT;
 RA Ikawa S., Tsuda Y., Fukada T., Sugimoto K., Himeno M.;
 RT "cDNA cloning of the Cytla receptor variants from Bombyx mori and
 CC their expression in mammalian cells.";
 RL Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: TO THE CADHERIN FAMILY.
 DR EMBL: AB041509; BAA99405.1; -;
 DR InterPro: IPR002126; Cadherin.
 DR Pfam: PF00028; cadherin; 6.
 DR SMART: SM00112; CA; 10.
 DR PROSITE: PS00232; CADHERIN_1; 1.
 DR PROSITE: PS00268; CADHERIN_2; 10.
 DR Calcium-binding; Cell adhesion; Glycoprotein.
 KW NON_TER 1716 1716
 SO SEQUENCE 1716 AA; 193323 MW; 0C2D5C3571097BC8 CRC64;

Query Match 62.8%; Score 5595.5; DB 5; Length 1716;
 Best Local Similarity 63.4%; Pred. No. 0;
 Matches 1093; Conservative 224; Mismatches 379; Indels 29; Gaps 12;

1 MGVEFFPAVLVLSAALANORCSYIAIPRETPELPIDYEGKMSQPLIPGPT 60
 1 MGVDRIIATLLIYAEVLAQRCGFMAIPRPRLPELDEGQWTWSORPLIPADR 60
 61 EGVCEMF---LPQMIQVIYMEEEIEGDVIAKLNYGSGNTPLV--SMSGOPRAQLCP 115
 61 EDVCGDYGHAMTPYGTGIIYMEEEIEGEVPIAKINRGPNVPIEPLFSSFLMLPV 120
 116 EFRONEADQMSLVITORDYETATMSTQSVSIQEGESQAVLVALEIVNIDNPILQV 175
 121 IRRIPDSNGEMHLLITQODYETPMQOYVFNIRIDGETLVAGVSLIIVNIDNAPIQA 180
 176 VSACVIRPHEGARLDYVYVSDRGELSTRPMTFRVDSRRADSIYMGVEYPSDMF 235
 181 LEPCCVDLGEARLRECYVVTYDADGRISTQPMQFRIDSDR-GDKKIFYIQANTPEMT 239
 236 NMKATVGINSPINFEETTLHIFSVYASDLSLNNHTVTMMVQYENESPPRAVEIFSYOO 295
 240 RMTVMVGINPELFTNPLHIFSVYALDSLPHTHVTLMVQYENEHPRRVEIFEAQQ 299
 296 FDEKTNQSFSLAIDGDTGNGINRAIYTLIRDDADPFSLVYE--DGAIIYVEIDRDK 352
 300 FDEKTAQSFPRALIDGDTGINKPIHRYLETAEDTFPHIRITEGSRGAILVDPIDRPT 359
 353 LERELFNLTIAVYKSTDSAPTEAHIFIVNDVNDORPEPLHKEYSIDIMEETPTLNPN 412
 360 LQREVFQSLITAIKIDNESSATIAANVAVIIVNDIDOREPEPLKEYRLNIMEETALTLPND 419
 413 EEEGFHDLGEMAOYTVLEDEVPFPGAASAFYIAPGSGYQORPTFMGTINTMLDYEDV 472
 420 QEEGFHDLGQNAQYTVLESDDYPADAKAIFYIAPEVGYQOTFMGTANMKMLDYEP 479
 473 IFQNIITIVYKADVNNASHVEGALYVNLINMNDLPTFEESYSASKEYVAGACFPVAT 532
 480 EFORIRLIVITIDMNEHGVAYVYVIMLNINDEEPIFEHSVQVNSFEKEGKGFVAN 539
 533 VLADRDIDDVVHSHNAVYDLFIDESTGEIFYSMDAPRYHONLTFYOVARDPTLG 592
 540 VRAHRDIDDVREHRLMGNANNYLSIDKDTGDIHVTDDEFYHNSQSELFYOVARDPTLG 599
 593 DGPANTVTLQVLELEDVNNPTPLRLPRSPSVSENVPEGEYIRELITATPPTDSAYLM 652
 600 E-PFTATISQLLIHDDIINNPTPLRLPRSPSVSENVPEGEYIRELITATPPTAEIR 658
 659 FEIDYTSYATKQGRAPNIEFNHCVEIETIYPAIINNGSAIGRLVVKKIRENVITDEE 718
 713 FEMLYLTVRVARDLNVIGDDDEAFITITIDMNDAPIFANGILTQIMRVELAAGTL 772
 719 FEMLYLTVRVARDLNVIGDDDEAFITITIDMNDNPPIWPGTLEOSLRKREMSDAGV 778

773 IGSVLTADIDGLYNOVRYTIOPRNNTEGLVKIDFTTGQIEVDANEAIDADEPWRFY 832
 779 IGTLTATIDIDGLYNOVRYTMYKANGETENLMDIFYTGQITVYKSGALIDADVPRRYNY 838
 833 YTVIASDECSLENRECEPDNSYFEVPGDIEIETIDTNKKVEPLTERKNTTVYMEANT 892
 839 YVVAVTDSRYADDPDCEPDPTMYETPGQVYIQTNNKIQPPEYDQFKAVVYIYEDAV 898
 893 SGDEYVQVLSHRDREDELHYRYTMMNFAVNRLDFFEVOLDTGRLEHYHNGDKLDRD 952
 899 SGDEVVAVVIGSLDRDDYHTIRYQINAVNRLDFFEVAVDPTGRVYVYITDDELDND 958
 953 GDEPTHTIFVNEIDNFFSDGGRNRQDEVEIFVLLDVNDNAPRLDELRFYDEGAV 1012
 959 GDEPHRIEFENLIDNFPQGGDGNRNQNDAEVLVLLDVNDNAPRLPEPDELMSVSESLT 1018
 1013 AGVRLPEIYAPDRDEPDTNSRVGIGLDTITTRDIEVPRDLFTMISTENKTELETPAM 1072
 1019 KGTRLQPHIYAPDRDEPDTNSRVGYAISLTIANRETEVEPLFTMIQIYNTGELTPAM 1078
 1073 DLRGVWGYEIPFIEAFDHDGYPOQSRNERYTLVIRPYNHHPPVFPPOPSVIRLSERAT 1132
 1079 DLRGVWGYEIPFIEAFDHDGYPOQSRNERYTLVIRPYNHHPPVFPPOPSVIRLSERAT 1138
 1133 EGVATATANELEPIYATDEDEGLHAGSVTFHVQNEAVOYFDTTEVGAGENSQILIR 1192
 1139 VAGLATATYDGEFLNRIYATDEDEGLHAGVAFEVGDTAEVDFYFHL--VNDGNSGTMLMK 1196
 1193 QLPPEQIRQFRITITATDGTGTEPGLMTDYTPSVYFVPIQGDPRVSENAATYAFTEGEG 1252
 1197 QLPPEQIRQFRITITATDGTGTEPGLMTDYTPSVYFVPIQGDPRVSENAATYAFTEGEG 1256
 1253 LRESFELPOAEDLKNHLEDDDCODIYYRFRIDGNNGELFLYDQSSNVISLAOLEBVAVS 1312
 1257 LERHELPRAEERKNNHLSDDCHNITYRKYVDGNNNGHGLDEITVNLVLYKLDVSYST 1316
 1313 YTLIAASNSPDATGIPLOTSLVYTVNVRANPRPIFEQDLYTAGISTLDSIGRELLTV 1372
 1317 YTLTIAASNSPTG-GIAL-TSTITTVNVRADPOPYFVRDLYTAGISTLDSINRELLT 1374
 1373 RASHREDDTITITIDRASQOLDSLEAVADSAFALHATTVGLVSLMKOPRASHNGHFEEDV 1432
 1375 QATSHENAPIYITIDLSVTWPTPLASVRETAFLNPHGVITLMIQTPASHNGHFEEDV 1434
 1433 IATPTASADTARVYVYLSSONRYTFIPDNQLETVQGRNFIAATEFSGFMNTNIDOV 1492
 1435 VATDPAAGSDRANVYKIIYLISTNRKVFLEFVNTLEQVEQNTDIEIAGTFSGFMNTNIDOV 1494
 1493 VFSDSGVAODDTTEVRAHPIRDVNVYQAOEVEAVRSQVTLKTIQMLSTNSVLQDL 1552
 1495 VPATDASGVIMNGITEVGRHFIIRDVNPVPADEIETLRGMVLLTGAIOSTLARTLVLRLD 1554
 1553 VTGDPPTLGEESMQTAVALALASVAGLCLVLLALFCRRALNRQOLASMTKYSY 1612
 1555 FT-DTSPAPDAGSAVALVALASVAGLCLVLLALFCRRALNRQOLASMTKYSY 1613
 1613 DSGLRNAGL-ARGTNKHAHEGSPMMNNEAIRAPDPAISDASGSDGLGIEDMPQFODY 1671
 1614 DSGLRNAGLARGTNKHAHEGSPMMNNEAIRAPDPAISDASGSDGLGIEDMPQFODY 1673
 1672 FPPGDTSSSGIVLHMGATDNKP--VTTGNNGFGKSTPYLPQ 1713
 1674 YPRRDV-----EFKTDKPEDYATANNNGFGKSTPSP 1708

RESULT 4
 Q9NDN3 PRELIMINARY; PR: 1716 AA.
 Q9NDN3
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE CADHERIN-LIKE MEMBRANE PROTEIN (FRAGMENT).
 GN B7175C.
 OS Bombyx mori (Silk moth).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditypsia;
 OC Bombycoidea; Bombycidae; Bombyx.
 NCBI_TaxId=7091;
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=MIDGUT;
 RA Ikawa S., Tsuda Y., Fukada T., Sugimoto K., Himeno M.,
 RT "cDNA cloning of the CryIa receptor variants from Bombyx mori and
 RL their expression in mammalian cells."
 Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO THE CADHERIN FAMILY.
 DR EMBU: AB041510; BAA99406.1; -.
 DR InterPro: IPR002126; Cadherin.
 DR Pfam: PF00028; cadherin; 6.
 SMART: SM00112; CA; 10.
 PROSITE: PS00232; CADHERIN_1; 1.
 DR PROSITE: PS50268; CADHERIN_2; 10.
 KM Calcium-binding; Cell adhesion; Glycoprotein.
 FT NON-TER 1716 1716
 SQ SEQUENCE 1716 AA; 193396 MW; 7A8151227C15C43A CRC64;

Query Match 62.8%; Score 5594.5; DB 5; Length 1716;
 Best Local Similarity 63.4%; Pred. No. 0;
 Matches 1093; Conservative 224; Mismatches 379; Indels 29; Gaps 12;

QY 1 MGEREFPALVLSLAAMLANORCSYIIAPRPPELPIDIECKSMSEOPPLIGPFR 60
 DB 1 MGVDAVILATLLILVETVAOERCGFMALPRPRDLPELDEGCTMSQRLLIAPDR 60
 QY 61 EBYCMENF---LPDMIOVYIMEEIEGDVYIAKLNYQSSNPVL--SIWSGPPRAQLG 115
 DB 61 EDVCMGYNHMPPTGTQIIMEEIEGEVPIAKLNYRGPNVPIEPAFLSGFNLLVPY 120
 QY 116 EEPQNRADGQMSLVITORDYETATMOSYVESIQVGEQANVLALEYINIDNPPILQY 175
 DB 121 IRIIPDSNGMHLITQKORDYETPGMOQYFNIRIDGETLVAGSVLLIYINIDNAPITQA 180
 QY 176 VSACVIREGEARLIDCVYQVSDRDEISTREMTFRVDSRADESIFYWGEYSDPMF 235
 DB 181 LERQYDELGEARLIECVYVVDADGRISTQFQGRIDSDR-GDKIIFYIGAGNIGEMT 239
 QY 236 NMAWYGINSPLEFETTOIHISVYASDCLPNNHYTMVQVBNESRPRRWYEIFSVQO 295
 DB 240 RMTMTGIMEPLFETNPPLHISVYALDSLPNHTVTLMVQVBNESRPRRWYEIFAVQO 299
 QY 296 FDEKTNQSLRADIDGDTGNRAINTLLIRDADDFEFLSEVIE--DGAIIHYTEIDRDK 352
 DB 300 FDEKTAQSPVRAIDGDTGINKIRHLETAEDYEFHRTLEGSGALIVYDIDRDT 359
 QY 353 LERLENLIVAKYSTDAFEEAHIFIIYNDVNDQRPRLKREYSIDIMEETPMLTNE 412
 DB 360 LQREYVQLSIKAYKNESATPAANYIIVNDIQRPRLKREYLNIMEETALLND 419
 QY 413 EEEGFHDLRDLGAOYVTELEDFEPGASAFYIAGSGYQRTFMGTINHTMLDYEDV 472
 DB 420 QEEGFHDLRDLGAOYVTELEDFEPGASAFYIAGSGYQRTFMGTINHTMLDYEDV 479
 QY 473 IFONTIIRKAVDMNNAHSVGEALVYVNLINNNDELPIFEESYSASEKETGAGPVPAT 532
 DB 480 EFORILARIADMDNEEYVAYIYINLINNDEEPIFEHSQNVSEKETGEGEVAN 539
 QY 533 VIALDRIDIDVYVHSLMGNAVYDLFIDESSTGELFVSMDDAFDYHRONTLFYOVARDTLG 592
 DB 540 VRAHNRIDIDRVREHITLMGANNVYLSIDKDTGDIHYQDODFPYHRQSELFYQVARDTLG 599
 QY 593 DGRHNVTITQVLEEDVNTPTLRLPRSTPSVEENVEGGEIENELIATDPDTSAYLM 652
 DB 600 E-PFHATISQLLIHDDINNTPTLRLPRGSPNVEENVEGYIITSEIRATDPDTTAEIR 658

QY 653 FEIDMDSTWATKOGRETNPEYGCIVETIYPEEGNRGSAIGRLVVOEIRDNVTIDEE 712
 DB 659 FEIDMTTSYATKOGREANPTEFNANCVETITYPAINNRGSAIRLVYKTRRENTVIDEE 718
 QY 713 FEMLYLTVRVDLNTVYIGDDYDEATFTITIDMNDNAPIFANGTLTQTMRELAASGTL 772
 DB 719 FEMLYLTVRVDLNTVYIGDDYDEATFTITIDMNDNAPIVWPGTLBQSLVREMSDAGVY 778
 QY 773 IGSVLATDIDGPELYNOVRTIOPRNNTPREGLVKIDTGTQIEVDANEAIDADEPRFYLX 832
 DB 779 IGLTATDIDGPELYNOVRTIOPRNNTPREGLVKIDTGTQIEVDANEAIDADEPRFYLX 838
 QY 833 YTVIADDESLERNKCPDPSNFYEPGDEIEIITDNNKVPPELREKFTYVYENAT 892
 DB 839 YTVIADDESLERNKCPDPSNFYEPGDEIEIITDNNKVPPELREKFTYVYENAT 898
 QY 893 SGDEVVQLYSHDRDELHYTVRYTNMFANVNPRLRDFEVDLDTGRLEHYHYPGDEKLRD 952
 DB 899 SGDEVVQLYSHDRDELHYTVRYTNMFANVNPRLRDFEVDLDTGRLEHYHYPGDEKLRD 958
 QY 953 GDEPTFTIYVNFIDNFFSDGGRNDEVEIEVYLLDVNDNAPPELPRFVSEGA 1012
 DB 959 GDEPTFTIYVNFIDNFFSDGGRNDEVEIEVYLLDVNDNAPPELPRFVSEGA 1018
 QY 1013 AGYRVLPETIAPDRDPRDNDNSRVYGIILDTITDDIDIVPOLFTMISTENKTELETAM 1072
 DB 1019 KGRRLDPHIYAPDRDPRDNDNSRVYGIILDTITDDIDIVPOLFTMISTENKTELETAM 1078
 QY 1073 DLRGYGTVEIEFLEAPDHGYPOORSNETYLVIRPNFHPVFPQPSVIRLSERAT 1132
 DB 1079 DLRGYGTVEIEFLEAPDHGYPOORSNETYLVIRPNFHPVFPQPSVIRLSERAT 1138
 QY 1133 EGGVLAATANELEPIYATDEDEGLHAGSVTFHVQNEAAYQVFDITEVAGENGOLILR 1192
 DB 1139 VNGILATVGEFENIRIYATDEDEGLHAGSVTFHVQNEAAYQVFDITEVAGENGOLILR 1196
 QY 1193 QLEPQIRQFRTIRATDGTGTEPMTDVTFSVYVPPGQDSEFSEMATYAFPEGE 1252
 DB 1197 QLEPQIRQFRTIRATDGTGTEPMTDVTFSVYVPPGQDSEFSEMATYAFPEGE 1256
 QY 1253 LRESFELPOAEDLKKNHLCEDDCDIYREFIDGNEGLFYVLDQSSNYISLAOELREVATS 1312
 DB 1257 LRESFELPOAEDLKKNHLCEDDCDIYREFIDGNEGLFYVLDQSSNYISLAOELREVATS 1316
 QY 1313 YTLHTAASNPDAATGIPLOTSLVTVVYVNRANPRIFEDDTYTAGISTLDSIGRELLTV 1372
 DB 1317 YTLHTAASNPDAATGIPLOTSLVTVVYVNRANPRIFEDDTYTAGISTLDSIGRELLTV 1374
 QY 1373 RASHTEDDTITYIDRASQDLSLEAVRDSAFALHATGVLISLNNQPPAASHMGPEFY 1432
 DB 1375 RASHTEDDTITYIDRASQDLSLEAVRDSAFALHATGVLISLNNQPPAASHMGPEFY 1434
 QY 1433 IATDPAASIDTARVKKVYLISQNRVYTFIDNQLETEYQNRNFIAATFSTGFNMTCNDIV 1492
 DB 1435 IATDPAASIDTARVKKVYLISQNRVYTFIDNQLETEYQNRNFIAATFSTGFNMTCNDIV 1494
 QY 1493 VEPSSSGVAAQDDTTEVRNHFTRDNVPVOAQEYEAIVRSVTYLTQMLSTNSLYLQDL 1552
 DB 1495 VEPSSSGVAAQDDTTEVRNHFTRDNVPVOAQEYEAIVRSVTYLTQMLSTNSLYLQDL 1554
 QY 1553 YNPDPTTGEESSQIATVYLAALASVIGLCVLLALALCFRRLANRLOALSMTKYGSV 1612
 DB 1555 YNPDPTTGEESSQIATVYLAALASVIGLCVLLALALCFRRLANRLOALSMTKYGSV 1613
 QY 1613 DGLNKRAGL-APGTNKAHAGSGNPMNMEAIRAPDEDAISDAGSDSLIGIEDMPQFDY 1671
 DB 1614 DGLNKRAGL-APGTNKAHAGSGNPMNMEAIRAPDEDAISDAGSDSLIGIEDMPQFDY 1673
 QY 1672 FPPGDTSSSGIYLVIMGEATDNKP---YTHHGNNGFSTPLPQ 1713
 DB 1674 FPPGDTSSSGIYLVIMGEATDNKP---YTHHGNNGFSTPLPQ 1718

RESULT 5
ID 09GRJ9 PRELIMINARY: PRT: 1717 AA.
AC 09GRJ9;
DT 01-MAR-2001 (TREMBLERL. 16, Created)
DT 01-MAR-2001 (TREMBLERL. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLERL. 17, Last annotation update)
DE CADHERIN-RELATED PROTEIN RECEPTOR BR-RI.
OC Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Sphingioidea; Sphingidae; Sphinginae; Manduca.
NCBI_Taxid=7130;
RN [1]
RP SEQUENCE FROM N.A.
RA Dorsch J.A., Maaly W.S.A., Griko N.B., Candas M., Bulla L.A., Jr.;
RT "A Cadherin-related Protein Receptor, BR-RI, in the Midgut Epithelium
of Manduca sexta Mediates Toxicity for Bacillus thuringiensis CryIa
Toxins."
RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RL -1- SIMILARITY: TO THE CADHERIN FAMILY.
EMBL, AF319973, AAC37912.1;
DN INTERPRO: IPR001525; C3_DNA_meth.
DR InterPro: IPR00126; Cadherin.
DR Pfam: PF00028; cadherin; 6.
DR SMART: SM00112; CA; 10.
DR PROSITE: PS00095; C5_MWASE_2; UNKNOWN_1.
DR PROSITE: PS00232; CADHERIN_1; 1.
DR PROSITE: PS00268; CADHERIN_2; 10.
DR Calcium-binding; Cell adhesion; Glycoprotein; Receptor.
KW SEQUENCE 1717 AA; 192305 MW; FEC4A48B098B17E CRC64;

Query Match 59.88; Score 5329.5; DB 5; Length 1117
Best Local Similarity 60.8; Pred. No. 0;
Matches 1049; Conservative 249; Mismatches 396; Indels 33; Gaps 16;

QY 1 MGVEFFPAVLIVSLASALANORCSYIIAIPR-PEPELPIDYEGKSMSEOPILPGPT 59
DB 1 MAVDVRIAPFLVFLAPLAVOERCGYMTAIPRLPRPNLPVLANEGOTWSGRPLPAPE 60
QY 60 REEVCMEFNE---LPDOMIQTVMEEIEEDVIAKLANQSGNTPLYST--MSGPRAOLG 114
DB 61 RDDCLMDAYHVTANGLQYIVLMEDEIDETITALNTYNGPSTPIELPFGSGVNL-LM 119
QY 115 PEFKONENDGOMSLVITRODIETATMOSYVPSIOVEGSOAVLVALEIVNIDNPILQ 174
DB 120 PYIRKVD-NGEMHLITIRORHYLEPQMOMYMNVRVDSQSLVAGVSLAIVNIDNAPIIQ 178
QY 175 VVSACVIEHGHARLTDCVVOYSDRDGEISTREPMFRVDSRAADESIFVMVGEVDPSPW 234
DB 179 NFEPCRVIELGPGELTCTYOVSADGRISTEFTFRIDSVR-GDEETFIERTINIPNMW 237
QY 235 FNMKATVGINSLPNETQTLHIFSVTASDSLPHNHTVTMNOVENVESRPWEVIFSVQ 294
DB 238 MVLNNTIGVNTSLNVTSPHIFSVTALDSLPHNHTVTMNOVENVESRPWEVIFSVQ 297
QY 295 QPEDTNSFSIRADIGTGIRKAINITLIRDDADFFSLEYI---EDGAILHVEIDRD 351
DB 298 QFEESYONFVIRADIGTEINMPINRYLTINEEDTFFSIEALPGKSGAVFLVSPID 357
QY 352 KLERLFNTIYAYKSTASPAFTEAHIFIVDVNDORREPLHKEYSIDIMEETPMTLNF 411
DB 358 TLQREVFLITVAYYDEFASTSTNVITVINDQREPRHKEYSIDIMEETPMTLNF 417
QY 412 NEEFGRHDDIGENQAYVELEVPEPPGASAFYIAPSGYQRTFIMGTINHTMLDYED 471
DB 418 DKEPFGHDKDIGNQAYVRLSVDPGAAEFYIAPEGYQRTFIMGTINHTMLDYED 477
QY 472 VIFONITIKKAVAMNASHGAEALVYNLNMWDELPIFEESSTYSASFKEVAGGFPVA 531
DB 478 PEFOSITIRVATDNNDRHVGVALLHIDLINMDEOPIFEHAQVOTVFDEDEGEGFVA 537

QY 532 TVLALDRIDVVVHSLGNADVLEFIDESTGEIFVSMDDAFDYHRQNTLFVORADDTL 591
DB 538 KAVANDRIGVVEHTLLGNVNFLLTKINGDIRVANSNDSPNRYRESELEFQVATDTL 597
QY 592 GDGPHNTVITQVLEEDVNTPTLRLPRSTPSVEENVPGCEISREITATDPTSL 651
DB 598 GE-PHNTSOLVIRLNDINNTPTLRLPRSGPQVEENVPGCHVTQELRATDPTTADL 656
QY 652 WEIDMDSTMTKQGRLENTPREYVGCIVERTIYPEGRGSAIGLVQOETRDNTIPE 711
DB 657 REEIMWDSFPAKQGRQANPDEFRNCVELTEFEPINNGLAIGVARELNHNTIIDE 716
QY 712 EEMLYLTVRVDLTNYIGDDDEATFTITIDMDNAPIFANGTLTQTMRYRELAASGT 771
DB 717 EREVSLTVRVDLTNYIGDDDESMITITIDMDNAPVWVEGTLQONFRREKASAGL 776
QY 772 LIGSVLATIDGPLYNOYRYTIQPPNNTPGELVKIDFTTGOIEVDANEAIDADEWRYL 831
DB 777 VVGSVRADDIDGPLYNOYRYTTFREDTDKDIMIDFLTGQISVNTSGAIDADPTPRHL 836
QY 832 YTVTASDECSLENTEBPPDSNFEVPGDIEIEIIDNNKVPEDTEKFNNTYVWENA 891
DB 837 YTVVADRCSTEDPADCPDPTTETEGNITITIDNNKVPQAEITKFDIVYIYENA 896
QY 892 TSGDEVOLYSHDRDDELVTYRTMFAVNPRLDEFEVDLTGRLEVHP---GDEK 948
DB 897 THLDEVVLIASDLDRDEIHTVSYVINYAVNPRLNFSVREGLVYVDYETQSGGEV 956
QY 949 LDRGDEPHHTITFVNFIDNESDGGRRQDEVELEFVLLVDYNDNAPEPLDELRFVYS 1008
DB 957 LDRDDEPHHTITFVNFIDNESDGGRRQDEVELEFVLLVDYNDNAPEPLDELRFVYS 1016
QY 1009 BGAVAGVRLPETYAPDRDEPTDNSRVGYGLDITLTDRODEVEPLDFTMISENKTGL 1068
DB 1017 ENLKQGVRLPETYAPDRDEPTDNSRVGYGLDITLTDRODEVEPLDFTMISENKTGL 1075
QY 1069 ETAMDLRGVYTEIFIEAFDGYPOORSNETYTLVIRPNHNFVFPDPSVIRLSR 1128
DB 1076 ETAMDLKQGWYTAHIAHFHGIPOMSNETYELIHFNHYAREFVPTDAVIRLAR 1135
QY 1129 ERATGCVLATANELEPIYATDEDCGLHAGSVTFHVQNEAAYQFDITTEVAGAGNSQ 1188
DB 1136 ERAVINGVATANGSEFLERISATDDEGLHAGSVTFHVQNEAAYQFDITTEVAGAGNSQ 1193
QY 1189 LILRLPFOIRKOPFLTRATDGTGTEPGPLMTVTFVSVEYFTQDGPVSENAATVAFPE 1248
DB 1194 LILRLDAPREIRKOPFLTRATDGTGTEPGPLMTVTFVSVEYFTQDGPVSENAATVAFPE 1253
QY 1249 GEEGRRESFELPOAEDLKHLCEDCODIYRFIDGNNGELFVLDOSSNVISLAOELRE 1308
DB 1254 KSAQMEHQPLADDKNHLCEDCODHSIYRIIDGNSGCHGLDPVNRRLKKELE 1313
QY 1309 VATSTLHTAASNSPDAGIPLOTLSLVTVVAREANPREPIEODLYRAGISTDSIGRE 1368
DB 1314 OSASHTLQVAAANSPPD-GIPLASITVTVVAREANPREPIEODLYRAGISTDSIGRE 1372
QY 1369 LITVRASTEDDTITVTDRASSMOLDSLEAVRDSAPALHATTVGLSLMOTYASHMGF 1428
DB 1373 LITVRASTEDDTITVTDRASSMOLDSLEAVRDSAPALHATTVGLSLMOTYASHMGF 1432
QY 1429 EEDVATIDTASADTARVAVYLIISSONRYTIFDNOLETVEONRNFIAATFSTGNNMCN 1488
DB 1433 KREVIATIDTAGQDRTDYVYVSSONRYYFVAVMTLQOVEDNRPFIDTFSAGNNMCN 1492
QY 1489 IDQVVPFSDS-SGVAODTTEVRAHFIRDNVVOQAEVAVNSDYVLTATOLMTSTNSL 1547
DB 1493 IDQVVPANPVYGVLEHSTQKRGHIFIRDNVYVLADEIOISDSVLTATOLMTSTNSL 1552
QY 1548 VLODVLGTPTPLGSESNQIAYYALASAVGFLCVILLALFQRTRLANQOLALSM 1607
DB 1553 VLODVLGTPTPLGSESNQIAYYALASAVGFLCVILLALFQRTRLANQOLALSM 1611

QY	1608	KYSVSDSGLNLRAGI-AGTKRHAVESNPMMNNAIAPOFDIASASGSDSLIGIEDMQ	1666
Dt	01-JAN-1998	(TREMBLrel. 05, last sequence update)	
Dt	01-JUN-2001	(TREMBLrel. 17, last annotation update)	
DE	INSECTICIDAL TOXIN RECEPTOR BT-R1 PRECURSOR.		
OS	Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).		
CC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;		
OX	Sphingioidea; Sphingidae; Sphinginae; Manduca.		
XX	NCBI_TaxID=7130;		
RN	[1]		
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.		
RC	TISSUE=MIDGUT;		
RX	MEDLINE=95197553; PubMed=7890666;		
RA	Vadlamudi R.K., Weber E.J., Li L., Ji T.H., Bulla L.A. Jr.,		
RT	"Cloning and expression of a receptor for an insecticidal toxin of		
RL	Bacillus thuringiensis."		
J	Biol. Chem. 270:5450-5454(1995).		
CC	-1- FUNCTION: BINDS TO THE CRYTA(B) TOXIN OF BACILLUS THURINGIENSIS		
CC	SUBSB. BERLINER, LEADING TO THE DEATH OF M. SEXTA.		
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.		
CC	-1- SIMILARITY: RELATED TO THE CADHERIN FAMILY OF CELL ADHESION		
CC	MOLECULES. CONTAINS 11 CADHERIN-TYPE REPEATS.		
DR	Interpro: IPRO01525; C5 DNA meth.		
DR	Interpro: IPRO02126; Cadherin.		
DR	Pfam: PF000128; cadherin; 5.		
DR	SMART: SM00112; CA; 10.		
DR	PROSITE: PS00095; C5_MTASE_2; UNKNOWN_1.		
DR	PROSITE: PS00232; CADHERIN_1; 1.		
DR	PROSITE: PS50268; CADHERIN_2; 10.		
KW	Receptor; Glycoprotein; Transmembrane; Signal; Repeat; Cell adhesion.		
FT	SIGNAL	1	21
FT	CHAIN	22	1528
FT	DOMAIN	22	1405
FT	TRANSMEM	72	1353
FT	DOMAIN	1406	1428
FT	DOMAIN	1429	1528
FT	REPEAT	72	116
FT	REPEAT	177	289
FT	REPEAT	290	397
FT	REPEAT	398	500
FT	REPEAT	501	623
FT	REPEAT	624	757
FT	REPEAT	758	882
FT	REPEAT	883	1004
FT	REPEAT	1005	1121
FT	REPEAT	1122	1242
FT	REPEAT	1243	1353
FT	CARBOHYD	127	127
FT	CARBOHYD	240	240
FT	CARBOHYD	246	246
FT	CARBOHYD	305	305
FT	CARBOHYD	468	468
FT	CARBOHYD	492	492
FT	CARBOHYD	575	575
FT	CARBOHYD	709	709
FT	CARBOHYD	865	865
FT	CARBOHYD	894	894
FT	CARBOHYD	986	986
FT	CARBOHYD	1049	1049
FT	CARBOHYD	1069	1069
DB	1672	EGFVEMDPVEYENKANG-----PEYANNNNNFAPNPFPSPPE	1708
QY	1667	FRRDYFPFGDDTSSTGLVMGEATDNKPETHGNNGFEKSTPYLPQ	1713
Db	1672	EGFVEMDPVEYENKANG-----PEYANNNNNFAPNPFPSPPE	1708
RESULT	6		
ID	P81137	PRELIMINARY;	PRT; 1528 AA.
AC	P81137:		
Dt	01-JAN-1998	(TREMBLrel. 05, Created)	
Dt	01-JUN-1998	(TREMBLrel. 05, last sequence update)	
Dt	01-JUN-2001	(TREMBLrel. 17, last annotation update)	
DE	INSECTICIDAL TOXIN RECEPTOR BT-R1 PRECURSOR.		
OS	Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).		
CC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;		
OX	Sphingioidea; Sphingidae; Sphinginae; Manduca.		
XX	NCBI_TaxID=7130;		
RN	[1]		
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.		
RC	TISSUE=MIDGUT;		
RX	MEDLINE=95197553; PubMed=7890666;		
RA	Vadlamudi R.K., Weber E.J., Li L., Ji T.H., Bulla L.A. Jr.,		
RT	"Cloning and expression of a receptor for an insecticidal toxin of		
RL	Bacillus thuringiensis."		
J	Biol. Chem. 270:5450-5454(1995).		
CC	-1- FUNCTION: BINDS TO THE CRYTA(B) TOXIN OF BACILLUS THURINGIENSIS		
CC	SUBSB. BERLINER, LEADING TO THE DEATH OF M. SEXTA.		
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.		
CC	-1- SIMILARITY: RELATED TO THE CADHERIN FAMILY OF CELL ADHESION		
CC	MOLECULES. CONTAINS 11 CADHERIN-TYPE REPEATS.		
DR	Interpro: IPRO01525; C5 DNA meth.		
DR	Interpro: IPRO02126; Cadherin.		
DR	Pfam: PF000128; cadherin; 5.		
DR	SMART: SM00112; CA; 10.		
DR	PROSITE: PS00095; C5_MTASE_2; UNKNOWN_1.		
DR	PROSITE: PS00232; CADHERIN_1; 1.		
DR	PROSITE: PS50268; CADHERIN_2; 10.		
KW	Receptor; Glycoprotein; Transmembrane; Signal; Repeat; Cell adhesion.		
FT	SIGNAL	1	21
FT	CHAIN	22	1528
FT	DOMAIN	22	1405
FT	TRANSMEM	72	1353
FT	DOMAIN	1406	1428
FT	DOMAIN	1429	1528
FT	REPEAT	72	116
FT	REPEAT	177	289
FT	REPEAT	290	3

FT	CAROHND	1104	1104	N-LINKED (GLCANC.) (POTENTIAL).
SEQ	SEQUENCE	1528	AA: 171986	MM: CF678E01D700C91D CRC64;
	Query Match	52.0%;	Score 4631;	DB 5; Length 1528;
	Best Local Similarity	60.0%;	Pred. No. 4e-287;	
	Matches	915;	Conservative 221;	Mismatches 366; Indels 22; Gaps 13;
QY	1 MGVRRFPAPVLLVSLASALANQCSYIIAIR-DEPELRPIDEGKSWSEQPLPGPT	59		
DB	1 MAVDRIAAFLFLVFAPAVLADQRCGYMTAIRPLRPRNLPALNEFGQTSWRPLPAPE	60		
QY	60 REVCQENF---LPQMIOVIYMEEDIEDVLIATLNTQGSTPYLSI--MSGRADLG	114		
DB	61 RDIIDMDAYHVTIATMIGIOVIYMEDEIDETIATLNTNGSPTEPLEPLFSLSYNL-LM	119		
QY	115 PEFRONADGGQSLVITPRODETATMOSYVSIOVEGSAQVLALEIVNTDNDPILQ	174		
DB	120 PYIRR--VDNGSASHHARQHTELPGMOQYMFNVAYDQGSILVAGSLAIIVNDNAPILQ	177		
QY	175 VVSACVIEHGEARLITDCVYQVSDRDGEISTFRMTFRYDSSRAADESIFYVGEYDPSDM	234		
DB	178 NFEPRVPELGGPLTECQYVSADGRISTEFMTFRIDSVR-GDEEFYIERTNIPIQW	236		
QY	235 FNMKATVGINSPLEFETQOLHIESYASDSLEPNHTVTMMQOVENVESRPPRWEIFSVQ	294		
DB	237 MVLNNTIGVNTSLNFTVSPHIFSTYALDSLPHTHTVMMQOVANNSRPRMELIFAVQ	296		
QY	295 QPDEKTNQSFSLRADISGPGINRAINYLIRPDADDFSLSEI--EGGALHVTIEDRD	351		
DB	297 QPEKSYNPFVRAIDGTEIIMPINRYLRITNEEDTFESIETLPGCKSGAFLVSPRIDD	356		
QY	352 KLERELFNLTIAVAKSTSPASFATEAHIFLIVANDVDQPEPLHKEYSIDIMEETPMTNF	411		
DB	357 TLQRRVPLTVIAKYKDEAPFSTGNVVIYVINDQREPIHKREYRIALIMEETPLTNF	416		
QY	412 NEEFGRHDBDLEGNQIYVELEDVPRPGSAASFYIAFGSGYQROFTINGTINHTMIDYED	471		
DB	417 DKEFGFHDKDQGNQIYVRLSEYDPRGAEEAFYIAPEGYQROFTINGTINHTMIDYED	476		
QY	472 VIFQNIITIKVAVDNNNSHGEALVYVNLIMWDELPIFEESYSASPKFEVGAQFPVA	531		
DB	477 PEFQGITIRVAVATDNDNRHGVALVHIDLINMDEQIFERAVQTVTFDEEGGCFEYA	536		
QY	532 TYLALDRIDIDVVVHSLGNNAVDFLEIDESTGELFVSMDDAFDYHRQNTLFOVARADPTL	591		
DB	537 KAVANDRIGDVEHTLLGNNAVNFLLDKLGDINVASDENSENRYRESELEFQVARTPTL	596		
QY	552 GDGPINVTTLQVLELEDVNNTPPTLRLPSTPSVEENVPEGEISREITATDPDTSAYL	651		
DB	557 GQ-PHRTISQVILRLANDINNTPTLRLPSTPSVEENVPAHVATITQELRATDPDPTADL	655		
QY	652 MPEIDMSTMTAKGQRENNPTREYVCCIVLEITTYPEGRGSAIGSLVVOELRDANTIDEE	711		
DB	656 RPELMDTSEFAKQGRQANPDEFRCVLEITTYPEINNRLGALIGVAVAREIRHNTIDYE	715		
QY	712 EFEMLYLTVARADLTVTYGDYDEAFPTITTIIDMDNAPIFANGLTQTMRELAASGT	771		
DB	716 EPEVSLTLVARRDLNTVYGDYDESMILLITTIIDMDNAPVWVEGLRBNFVREMSAGL	775		
QY	772 LIGSVLATIDGPLYNOVRYTIQPRNNTPPEGVLKIDFTTGQIEVDANADADAEFWRYL	831		
DB	776 VVGSVRAADIDGPLYNOVRYTIPEPREDYDKDLMIELPHGSAFNRHKKRIDANTPPREHL	835		
QY	832 YTVTAVASDECSLENNTEPCPPDSNFEVPGDIEIEIDTNNNKVPBELTEKKFMTYVVMENA	891		
DB	836 YTVTAVASRCESTEDADCPDPPTYMETEGNITITHTDNNKVPQAEETKPTVYIYENA	895		
QY	892 TSGDEVOLYLSHDBRDRLTVYRTYTNMFANPNRLRDFEVDLTQGRLEAVP--GDEK	948		
DB	896 THLDEVYVLIASDLDRLDILYHMSYVIYANPNRLMNFVSNNREGLYVYDYETQSGSEV	955		
QY	949 LDRDDEPTHTLTVNFIDNFTSDGGRNNOVELEIVVLLDVNDNAPMLPDELRFDSV	1008		

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2002, 13:34:02 ; Search time 7575.83 Seconds

(without alignments)
11972.483 Million cell updates/sec

Title: US-09-715-909-1
Perfect score: 5498
Sequence: 1 cataatacataaagaaga.....aaaaaaaaaaactcgag 5498

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues
1 number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vl:*
30: em_hcgo_hum:*
31: em_hcgo_inv:*
32: em_hcgo_rod:*
33: em_hcgo_hum:*
34: em_hcgo_inv:*
35: em_hcgo_rod:*
36: em_hcgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	5498	100	5498	6	AX147201	AX147201 Sequence
2	20178	36.7	5582	3	AF319973	AF319973 Manduca s
3	1982	36.0	5469	3	AB026260	AB026260 Bombyx mo
4	1975.2	35.9	5172	3	AB041508	AB041508 Bombyx mo
5	1964	35.7	5172	3	AB041510	AB041510 Bombyx mo
6	1959.2	35.6	5172	3	AB041509	AB041509 Bombyx mo
7	1954.6	35.6	5577	6	AR096316	AR096316 Sequence
8	1954.6	35.6	5577	6	AR096316	AR096316 Sequence
9	1946	35.4	5355	3	AF367362	AF367362 Heliothis
10	1826.6	33.2	5527	6	AX147203	AX147203 Sequence
11	1736	31.6	5527	6	AX150183	AX150183 Sequence
12	1547.4	28.1	5592	6	AX147207	AX147207 Sequence
13	1465.8	26.7	5592	6	AX147205	AX147205 Sequence
14	1363.2	24.8	7799	3	AE367363	AE367363 Heliothis
15	61.8	1.1	11419	1	AE004760	AE004760 Pseudomon
16	58.4	1.1	52863	2	AC017834	AC017834 Drosophill
17	58.4	1.1	174500	3	AC009202	AC009202 Drosophill
18	58.4	1.1	251154	3	AE003656	AE003656 Drosophill
19	58.4	1.1	294942	3	AC011662	AC011662 Drosophill
20	57.2	1.0	10087	1	AE005045	AE005045 Halobacte
21	56	1.0	35856	1	SC59	AL138851 Streptomy
22	55.2	1.0	37931	1	SCD10	AL135988 Streptomy
23	55	1.0	1319	1	STMIMP	D00670 Streptomyce
24	55	1.0	1899	1	STMIMP	M89476 Streptomyce
25	55	1.0	2007	1	AB042262	AB042262 Corynebac
26	55	1.0	2045	1	STMENDOPRO	M81703 Streptomyce
27	55	1.0	2261	6	I20933	I20933 Sequence 1
28	55	1.0	41807	1	SC6D11	AL158061 Streptomy
29	54.6	1.0	196490	2	AC005507	AC005507 Plasmodiu
30	54.2	1.0	3051	1	ABU20583	U20583 Azospirillu
31	54	1.0	13040	1	AE005769	AE005769 Caulobact
32	54	1.0	152973	2	AP003334	AP003334 Oryza sat
33	53.4	1.0	24800	1	SC9B1	AL049727 Streptomy
34	53	1.0	22775	1	SC5G8	AL33872 Streptomy
35	53	1.0	36144	1	SC9C5	AL357523 Streptomy
36	53	1.0	113193	1	AF357202	AF357202 Streptomy
37	52.8	1.0	10982	1	AE005834	AE005834 Caulobact
38	52.6	1.0	123580	1	AF263912	AF263912 Streptomy
39	52.4	1.0	37123	1	SC5F7	AE006872 Streptomy
40	52.2	0.9	11623	1	AE002066	AE002066 Deinococc
41	52.2	0.9	12895	1	AE008084	AE008084 Agrobacte
42	52.2	0.9	14860	1	AE007093	AE007093 Mycobacte
43	52.2	0.9	63033	1	MTV008	AL021246 Mycobacte
44	52.2	0.9	199551	2	AC006281	AC006281 Plasmodiu
45	52	0.9	1141	8	HVU26545	U26545 Hordeum vul

ALIGNMENTS

RESULT 1
LOCUS AX147201 5498 bp DNA
DEFINITION Sequence 1 from Patent WO0136639.
ACCESSION AX147201
VERSION AX147201.1 GI:14346372
KEYWORDS
SOURCE
ORGANISM
European corn borer.
Ostrinia nubilalis

REFERENCE
AUTHORS
TITLE
JOURNAL
PIONEER HI-BRED INTERNATIONAL, INC. (US)
Location/Qualifiers
1. 5498
/organism="Ostrinia nubilalis"

FEATURES
source
PAT 08-JUN-2001
Applicant

CDS

/db_xref="taxon:29057"
162..5315
/note="unnamed protein product"
/protein_id="CAC41165.1"
/db_xref="GI:14346373"
/translation="MGVERFPAPLVLSAALANORCSYIIAIPRETELPIIDY
EGKMSDPLIPGPTREVECMENLPPDMIOVITMEBEDVIAIKANOGSMPLV
SINGORPAROLGPEFRONADGMSLVITQODYETARMOSYVSIOVEGSOAVIA
LEIVNIDNPILIDVSAVILPEHGEALIDCVVDSRDEISITRETFHVSRAA
DESIFVNGEYDPSDMFNKMTVGINSPLNETQHLFEVYASDSLEPNHTVIMVO
VENVESRPRMWEIFSVQODEKTNOSLSAIDGDIINAIYTLIRDADDEFSL
EVIDGALHYTEIDRDLERLEENLTIVAKSDASPAFAHFEIYNDNDORPEP
LHKYEYSLIDMETPMPLNENFEFGFHDNRDENNOYVELDEPGGAASAFAPGS
GYOQRTFMCTMINTMDYEDVIRONTIIRKAYADMNNAISVGLAYVNIIMNDEL
PIEESYSASFETVAGFPVATVALDRIDIVYVHSLGNADVILFIDESGEIT
VSMODAFYHRQNTLFEVQVRADPLGDSGPHMTVTQVLELEDVNNPPLRLRSTP
SVENBVEGEYISREITATDPDTSAYLMFEIDMSYTWAKGRETNPTEYGCIVET
IYPEGNRGSAIGRLVVOEIRDNVITDEFEEMLYLVVRDLMTVIGDDYDEATFTI
TIIDMNAPLEFANGTLQTRVRELAASGLISLVATLIDGFLAYNOVRTIOPRN
TPBEGLKIDFTTGOLEVDANEAIDADEPMRYLYVTYASDECSLENTCEPDSNF
EVPEDTIEITIDTNKKPEPELEKENTVYWEAATSGDEVOLXSHDRNDELHYH
RTINFAVNPRKDFEVDLDTGRLEVHPEDENLDRGDEPHTIFNFDNFSDS
DGRNODEVEIFEVLLVDVNDAPBMPLEDELFRVSEGAAGVAVLEPIYAPDEPD
TDSRVGIGLIDLTITDRDIEVPLFTMISIENTKGETEMDRGVTGEIETFEAF
DHGYPQORSNETYLVIRYFNHFPVFPQDSVIRLSRERATMGGLATANAMELE
PIYATDGGELAGSVTFHVOGNEAAYOFDITEVAGAGNSGLILROLFPRIOPRI
TIRADGTERGRLMTDVTESVFPVPGDSEPMATVAFPEGEBSLRESPLPAS
EDLNHLCEDDCODIYTRFDIGNEGFLYLDOSNNVLSLADELREAVTSTTLIAS
NSPATGIPLOTSLIVTVNREANPRITFEODLYTAGISLUDISIGRELTVRASHE
DDTITVYIDRASQDLSLEAVRDSAFALHAATGVLISLMQPTASMHGMEFEDVATD
TASAIIDYARVAVYLISQNRVTFEFDNOLEVEQNRMFIATFSTGFMTCNIDQVP
FSDSGAODDETEVRAHFIRDNVVOAEVAVRSPLELRTIOLMSTSLYDLDL
VTGPTPLIGESMOIAYALALASVIGFLEVLILFCGTRALNRDLOLMSMKYG
SVDSGLNRAGLAPSTNKHAVGSPNMEAIRADPDALISASDSDLIGEDMPOFR
DDIFPEGDTSSSIVLHMGFATDNKPYTTHGNMFKESTYLPQPHRK"

BASE COUNT 1428 a 1537 c 1434 g 1099 t
ORIGIN

Query Match 100.0%; Score 5498; DB 6; Length 5498;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 cataatacaataaagaagatgtgttgaataaagaagaagttaataaacttgataa 60
DB 1 CATATAATACAAATAAGAGAACTGTGTGAAAAAGAAAGATTATTAACCTGGATTA 60
61 ttaaaccttgaaaaaacccggtgtttaagtggaaattttgtctgaagacaacccgttggata 120
DB 1 TTAACCTGAAAAAAACCGGTGTTTAAGTGAATTTTGTCTGAAGACAAACCGTGGATA 120
QY 121 gctcaaatataaatttactactaagaatacatgcaaaatgggggtttggaagttct 180
DB 121 GCTCAAAATTAATAAATTTCTACTAATACTAAGATCATGCAAAATGGGGGTGAGAGTTCT 180
QY 181 tcccagcagctgactggtctctctttagcctctgcgcgactaagccaaccaacatgttcgt 240
DB 181 TCCAGCAGGTGCTACTGTCCTTTAGCCTCGCCGCACTACCAACCAACGATTTGCT 240
QY 241 acattatcgcaatacaagaaccggaactccggaactgcgcctatctgattacgaagaa 300
DB 241 ACATTATCGCAATACCAAGACCGGAGACTCCGGAACCTCCGCTATTGTACGAAGANA 300
QY 301 aatactgagtgaaacagctcttaatacccgccggaaccggagaagaaatgtatgtggaga 360
DB 301 AATCATGAGAGTGAACAGCCTTAATAACCCGGCCGACCCGAGAGAGAAAGATGTATGAGAA 360
QY 361 acttttaaccggaataaattgattcagatcatatataggaagaagaataacgaagtagag 420
DB 361 ACTTTTACCGGATTAATAATGATTGAGGTCATATACATGGAGGAANAATCGAAGAGACG 420
QY 421 tcatcatgtcggaagcttaactatacaagggtccaacacgcccgtgtgtcogattatgcag 480

|||||
DB 421 TCATCATTCGGAAGCTTAATCAATGAGGTCACACACGCGGCTGCTGCATGATATGTCAG 480
QY 481 gccagcccaagagcccaagcttggtgccccttagtttcgacagaaatgaagcagaacggtccaa 540
DB 481 GCCAGCCCAAGAGCCCAAGCTGGGGCCCTGAGTTTCGACAAATGCAACACGCGCAATGGA 540
QY 541 gctctgtattacgcaaaagacaagaactcagagacagcaaccaatgagagcattgtctct 600
DB 541 GCTTGTTATTATGCAAAAGACAAAGACTACGAGACGCAACCATGCAAGACTATGTGTTCT 600
QY 601 caatcaagctggaaggtgaatcacagagccgtactgtgtgccttggaagaaatgtaacaaatcg 660
DB 601 CAATCAAGTGAAGGTTGATATCAGAGCCGCTACTGCTGGCGCTGAGATGATCAACATCG 660
QY 661 acgacaatccgcccactcctctgcaagtgtgtcagcgctgttaattccagaacatgycagag 720
DB 661 ACGACAATCCGCCCACATCTCGCAAGTGGTCAAGCCTGGTAAVTCAGAAACATGCGCAGG 720
QY 721 ctgagctgacgactgctgtatcccaagttgtcagaccgcaagtgaaaaatcaagaccgct 780
DB 721 CTAGACTGACGACTGCGGTGACCAAGTGTACAGACCGGACGCTGAATACAGACCCGCT 780
QY 781 tcatgacgttccggtgtcgacagcagcagggctgcagatgaagaacatcttctacatgtgtg 840
DB 781 TCATGACGTTCCGTGTCGACACAGAGGCTGCAGATGAAGCAATCTTCTCATAGTTTG 840
QY 841 gagaatagaccgcccgagctgtgtcaatatgaatgactgtggagatcaatttgcct 900
DB 841 GAGAAATAGACCCCGACGACTGTGTTCAATATGAATGACTGTGGGATCAATTTGGCCT 900
QY 901 tgaacttcgaagaacactcagcttcataatattagcgtcacagcttcctcgtaccga 960
DB 901 TGAATTCGAGCAACTCAGCTTCATATATTAGGCTCACACCTTCTGACTGCTACCGA 960
QY 961 acaacacaacggtcaacatgatgtgtcagaagtggagaacgttagagctccggcccctgct 1020
DB 961 ACAACACACACGGTCAACCATGATGTCGAAGTGGAGAAGCTAGAGTCTGGCCCCCTCGCT 1020
QY 1021 ggggtgagatctctcagatgacgaggtttgagagaagaacataagaagcttctccccc 1080
DB 1021 GGGTGAGATCTTCTCAGTGACGACGAGTTTGACGAAGAAGCTATATAGAGCTTCTCCCTCC 1080
QY 1081 ggcgcataagcgggagacacgggaatacaatagggccatcaatataccctcatcaggagtg 1140
DB 1081 GCGGATAGAGGGGACACGGGAATCAATAGGGCATCAACTATACCCATCAGAGGATG 1140
QY 1141 acgctgaacacttcttcccttgaggtgatatgaaagcaggagctattctgcagtgactg 1200
DB 1141 ACGCTGACGACTTCTTTCCCTGGAGGGTATTGAAGACGAGACTTCTGACAGTGACTG 1200
QY 1201 agatgaccgcgacaagaacttgaagaagagctttcaacttccacatctgttcttcaaat 1260
DB 1201 AGATGACCGCGACAAAGCTTGAAAGAAGACTTTCAACCTCCACATCGTGTTCATTAAT 1260
QY 1261 ctactgaagctagcttctgcaacagagcccaactttcaatctcgtlcaacgaagctcaatg 1320
DB 1261 CTACTGAGCGTACTTGTGCAACGAGGCCCAACATTTTCATCTGTCACAGAGCTCAATG 1320
QY 1321 atcagcagaccgagccgctgtgcaataaagaatacagtatgatatacatgtgaggaactcaa 1380
DB 1321 ATCAGCAGCCGAGCCGCTGCTAATAAAGATACAGTATGATTCATGAGAGGAACCTCAA 1380
QY 1381 tgaacttaacttaataaagaatttgatccaatgactcagatcgagatttgggtgaaacgctc 1440
DB 1381 TGACTCTTAACCTCAATGAAGAATTTGGATTCATGATCGAATTTGGGTGAAGAGCTC 1440
QY 1441 aatacagctggaacttgaagagctgttccgccaggggqggctgcgcacatctcatcag 1500
DB 1441 AATACAGTGAATCTTGAGAGAGCTTCCGCCAGGGGCGCTCCGATTTCTACATCG 1500
QY 1501 cgcgggagcggctacacagaggcagaccttcaatcatatgggcaacataaacacacatgac 1560

Db 1501 CCCCGGGAGCGCTACAGAGAGACCTTCATCATGGGACCACTAAACCAACCATGC 1560
QY 1561 tggattagaagatgtcatcttttcagaacatcatatcaagtcgaagcagtgagacata 1620
Db 1561 tggattagaagatgtcatcttttcagaacatcatatcaagtcgaagcagtgagacata 1620
QY 1621 acaagcgaagcagtgagcagcagcgtgtgtacgtgaactgtatcaactgtgaagcagc 1680
Db 1621 ACAAGGCTAGGACGAGGCGGAGGCGCTGTACGTGAACCTGATCAACTGGAAGAGAG 1680
QY 1681 aacttccactcttcgaagagagcagctactccgcgtcgtttaaggagacgcgtgcgcg 1740
Db 1681 AACTTCCACTCTTCGAGAGAGACGACTACTCCGCGCTTAAAGAGAGACGCTGGGCGCG 1740
QY 1741 gtttcccggtgagcagcagtgctgcgcgtcccttcgaagagacatgaagcagtaagtgtat 1800
Db 1741 GCTTCCCGGTGAGCGAGGCGCTGCGCTCGAGAGACATCGACGACGCTGATGTCATT 1800
QY 1801 catgtatggcagcagctgtgtactaccgtgtcatatagatgaatcaacgggagagatcttcg 1860
Db 1801 CATGTATGGGCAACGCTGTGTGACTACTCTTCAATGATGAATCAACGGGAGATCTTTCG 1860
QY 1861 tgaagatgagcagatgctctcgaactacccaagacagacatctatctgttcagctgtgcg 1920
Db 1861 TGAGCATGAGCATGCGCTTCGACCTACACCGACAGACACTCTATTGTTCAGGTGCGCG 1920
QY 1921 ctgaagcagcttctgtgagcagcagccacacacacagtgacacccaagctgtgtatagaac 1980
Db 1921 CTGACGATACCTTGTGGCGGCGGCGCACACACAGTGCACCCAGCTGTGATGTAAC 1980
QY 1981 tggagatgtcaacaacactcctccacacactcgccttgcgcggttcgaactcgaagcgttcg 2040
Db 1981 TGAGAGATGTCAACAACACTCTCTCCACCTTACGCTTCCGCGCTTCGACTCCAAAGCTTCG 2040
QY 2041 aggagacagcttcgcaagagatacagagataccgggaaatacactgtctacagacccggaca 2100
Db 2041 AGGAAAGAGTTCCTCCAAAGATACGAGATATCCGGGAAATCACTGCTACCGACCGGACA 2100
QY 2101 ccaagcctactcgtgtgtgagatcgaactggaactcaacctgggacccaagcagcgga 2160
Db 2101 CCAAGGCTACTCTGTGTGATCGAGATCGAGTGGACTCCACTGGGCGCACCAAGCGAGGA 2160
QY 2161 gagagaccaaactactaataacgtcgggtgtatagttatcgaacagatatccccacgc 2220
Db 2161 GAGAAACCAACCTCTACTGATGATGCGGGGTATAGTTATCCAAACGATATACCCACCG 2220
QY 2221 agggcaacgggttcgcgcacatcgggcgcctcgtgtgtcgaagagataccgggacaaactca 2280
Db 2221 AGGGCAACCGGGGTTCCGCGCATCGGGCGCCTGTGTGTCGAAGAGATCCGGGACAACTGCA 2280
QY 2281 ccatgacttcgagaatctcgagatgtcttactcaactcgcgttcggtgtggagacccaca 2340
Db 2281 CCATGACTTCGAGAAATTCGAGATGCTTTTACCTACCGTCCGCTGAGGAGCTCAACA 2340
QY 2341 ctgtcatcggagatgaactaagatgagcgaagcttcaagatcaacaataatcgacatgaagc 2400
Db 2341 CTGTCTATCGGAATGACTATGATGAGGCGAGCTTCAAGATCAATTAATCGACATGAAGC 2400
QY 2401 acaacgcgcagactcttcgcaagcagcagctgtgcgcgaagagatgtgcgtgcgcagcgtgcg 2460
Db 2401 ACMAAGCGCGGATCTTCCGAAACGCGACGCTGACGACAGATCGGCTGCGCGAGCTGG 2460
QY 2461 cggcgaagcgcagcgtctcgtcgtcgtcgtcgcgcgaagacatgaagcccgctctca 2520
Db 2461 CGGCGAGCGGACGCTCATCGGCTCCGCTCGGCTCGCAACGCACTCGACGCGCCGCTCTACA 2520
QY 2521 accaagtgcgtacatactacacactagaacaacacactcccgaggatgaatgaatgtg 2580
Db 2521 ACCAAGTGGGTACTACTATACAACTAGAAACAAACACTCCGAGGAGTTAAGTAAGATTG 2580
QY 2581 acttcaacactgtgtcaaatgtgagtgagtgcaacgagcgatcgaatgcagaagaaacct 2640
Db 2581 ACTTCAACACTGTGTCAAAATGTGAGTGTGGAACGAGGCGATCGATCGACGAAACCT 2640

QY 2641 ggcgcttctactgttactacacgcgtcatcgtcragcagagtggtccctcgaaaccgca 2700
Db 2641 GCGCTTCTACTTGTACTACACGCTGATCGTACGACAGTGTCTCCGTGAAACCGCA 2700
QY 2701 cggaaatgtccctcagaatccaaactactcgaagttccaaagcgatcgaatagaata 2760
Db 2701 CGGAATGTCTCCAGATTTCCAACTACTTGGAAAGTTCCAGCGCATATCGAAATAGAATCA 2760
QY 2761 tcgacaacaacaagaatgtcctgagcgcgtcaactgtagaagttcaacacgagctgtacg 2820
Db 2761 TCGACAAACAACAAGAGCTGTGAGCCCTCATGTGAGAAAGTTCAACAGAGAGGTGTACG 2820
QY 2821 tctggagagatgccaacgagcgcgcgaagctgttcgaagtgtaacttcccaagccgtgaca 2880
Db 2821 TCTGGAGAGATCCACAGACGGCGGACGAGAGTGTCTCAGCTGTATCTCCACGACCGTACA 2880
QY 2881 ggaacgagctgttaccacaacagtgatagatacagatgaacttgcgtgtgaaccccgactgc 2940
Db 2881 GAGAGAGCTGTACCAACAGGTAAGATACGATACACTTTCGGGTGAACCCCGACTGC 2940
QY 2941 gggattcttcgaagtgagacctgtgaacactgtgcgcttgaagtgatcaaccggggagc 3000
Db 2941 GGGATTTCTTCAGAGTGGACCTGTGACACTGTGCGCTTGAAGGTGATACCCGGGGAGC 3000
QY 3001 aaaaattggaacgcgagtgagtgagctacacatactatcttgaatcaatcgcgata 3060
Db 3001 AAAAAATTGAGCCGATGGGGATGAGCCCTACACATATATCTTTGTAATTCATCTGATA 3060
QY 3061 acttcttctcgtatgtgtacagtgagagaaacagagacgaagtgaatattgttcgttc 3120
Db 3061 ACTTCTTTCTGTATGTGTACGCTAGGAGAAACAGAGAACTTGAATATTGTTCGTTTC 3120
QY 3121 tatgtatgtgaacgaacagcgtcctgaagatgcatltgcctgaatgaactcggtttga 3180
Db 3121 TATGTGATGTGAACGACAACGCTCTGAGATGCATTGCGTATGATCACTCCGTTTGAATG 3180
QY 3181 ttccgaagagcagatgtgtgtgtcgtgtgtactcccaagaaatcagcagccggagcag 3240
Db 3181 TTTCCGAAGAGACCAATGTGTGTGTCCGTGTCTCCGAAATCTACGACCGGACAGG 3240
QY 3241 atgaacgaacagcagacaaactcgtgtcgtgtgaagaaatcccttgacactcaacgacg 3300
Db 3241 ATGAACCAAGACAGGACAACCTGCGTGTGCTGATGGAATCTTGGAACCTCAAGATCACCG 3300
QY 3301 accgagacatcgaagtgtccggaatctctcaaccaatgtcgaatgaatgaatgaatgaat 3360
Db 3301 ACCGAGACATCGAGGTGCGCGGATCTCTTCAACATGATCTCGATTGAAAAACAACCTGGG 3360
QY 3361 aacttgaacgcgctatgtgacttgaaggggtatgtggcaacttgaagaatatcatatgag 3420
Db 3361 AACTTGAAGACCGCTATGAGCTTGAAGGGGTATTTGGGCACTTACCAAAATATTCATTGAGG 3420
QY 3421 ccttcgaacagcagctacccgagcagaggtcccaagagctacacccctgtgcacccgc 3480
Db 3421 CCTTGACACGCGCTACCCGAGAGAGTCCAAAGAACGTAACCTGTGATCTCCGCC 3480
QY 3481 cctacaacttccacacccctgtgtcgtgttcccgcaaccccgactcgtlcatctggtctt 3540
Db 3481 CCTACAACCTTCCACACCCCTGTGTGTGTCCGCAACCCGACTCCGATTCGGCTTT 3540
QY 3541 ctgaagagcgcgcaacaaagagcgcgtcttcgtgcgaagcgttcgaagagatctccggagc 3600
Db 3541 CTGAGGAGCGGCAACAAAGGCGGCTTGTGGCCAGCGCTGCCAAACGATTCCTGTGAGC 3600
QY 3601 cgaatcagccacagcagagagcgcgtcccaagcgcgcgcgttcacgttccacgtccag 3660
Db 3601 CGATCTAGGCCACGAGAGAGAGGCGCTCCACGCGGAGGCTTCCAGCTCCAGG 3660
QY 3661 gaaatgagagcgcgttcaagtaacttgaataactgaagtgtggagcagagagaaatagcg 3720
Db 3661 GAAATGAGAGAGCGCTTCAATCTTGTATTAACGAAGTGGAGACAGAGAAATATACG 3720

QY 3721 ggcagctatataatcgcagctcttcccaagagcaaatcagaacaaatcagatcagatcc 3780
 |||||||
 Db 3721 GGCAGCTATATATTCAGCGAGCTTTCCAGAGCAAAATCAGAAATTCAGATCCAGATCC 3780
 |||||||
 QY 3781 gggcgaagagagcgagcgagcgccgctttgagccgaagatcagtttccgtgag 3840
 |||||||
 Db 3781 GGGCGACGAGGAGGCGGAGCGGCCCGCTTTGGACCGAGTCACGTTTTCGGTGG 3840
 |||||||
 QY 3841 tcttcgtaccccaacagggcgaccagtggttcagcgaaagcgagctctctcgtctct 3900
 |||||||
 Db 3841 TCTTCGTACCCACACAGGCGGACCCAGTGTTCACGAAAAATGACGCTCTCTCCCTTCT 3900
 |||||||
 QY 3901 tccagggatgaagagggcctccgttagagatttgagctgcgcaagcagaagacctaa 3960
 |||||||
 Db 3901 TCGAGGGTGAAGAGGCGCTCCGTGAGAGTTTGGAGCTGCCCAAGCAAGAACCTTTAAA 3960
 |||||||
 QY 3961 accacccctcgaagagagctgcgcaagatatctactacagttatttgacggcaacag 4020
 |||||||
 Db 3961 ACCACCTCTGGAAGATACGCAAGATATCTACTACAGGTTTATGACGGCAACACG 4020
 |||||||
 QY 4021 agggctcttctgtaactgagcagtcgaacgcaacgtcatctcccttcgaggaagtga 4080
 |||||||
 Db 4021 AGGGCTCTTTGCTACTGAGACCAAGTCACGATCATCTCCCTTCCGACGAGATTTGAC 4080
 |||||||
 QY 4081 gcgaggtgagcagctctacacgctgcacatcgcgagcaactgcgcgagccactg 4140
 |||||||
 Db 4081 GCGAGGTGCGCACGCTTACACGCTGCACATCGCGGAGCAATCGGCCACGCGCATCTG 4140
 |||||||
 QY 4141 ggtatccctctgagactctcctcgttgcagagtgtaagtgaagaagcgaaacggc 4200
 |||||||
 Db 4141 GGATCCCTCTCAGACTTCCTCCTGTTGACGGTGAATGTAAGAAAGAAAGAACCCGC 4200
 |||||||
 QY 4201 gcccaattctcgaagcagcactctacacagcgagcgatctcagcttgagcagcatg 4260
 |||||||
 Db 4201 GCCCAATTTTGAGAGAGACCTTTACACAGCGGGCATTTTCAGGTTGACAGCATTTGAC 4260
 |||||||
 QY 4261 ggggaattgctactgctcagggcgagcgacacagaagaagacacacacacacacac 4320
 |||||||
 Db 4261 GGGGAATTTGCTTACTGTCAGGGGAGCCACACAGAAAGAACACCATCTACACCATAG 4320
 |||||||
 QY 4321 accgtcgagatcagctgagctgacagcagctagaagcgctgcgcaactcgccctcg 4380
 |||||||
 Db 4321 ACCGTGCAAGATGAGCTGAGACAGCCTTAGAAGCCGTGCGCATCTCGGCTTCGGGC 4380
 |||||||
 QY 4381 tgcatacgacacacggcgctgtcttcgctcaatatacgacccacggcttcacatgca 4440
 |||||||
 Db 4381 TGCATGCGACACCGGGCTGCTTTCGCTCAATATGACACCCACGCTTCATGACGGCA 4440
 |||||||
 QY 4441 tgttcgagtgcaagctacgctacgatacagctctgcaatcgacacagccgctgta 4500
 |||||||
 Db 4441 TGTTCGAGTTGACGCTACCTTACGCTTACGATGAGTTCGCAATCGACACAGCCCTGTA 4500
 |||||||
 QY 4501 aagctactcatctcatcgaacaaacggcgagctctcatcttgataacacactgga 4560
 |||||||
 Db 4501 AAGTCTACTCATCTCATCGCAAAAACCGCTGACCTTATTTTCATTAACCAACTTGA 4560
 |||||||
 QY 4561 ccgttgagcagaagaatttcaagcgccagctcagacccgggttcaacatgaagt 4620
 |||||||
 Db 4561 CCGTTGACACAGAAATTTCAATAGCGGCCACCTTCAGACCCGGGTTCAACATGAGCT 4620
 |||||||
 QY 4621 gcaacatcgacacagtggtgctgcgttcagcgacagcagcggtgagcagaacacaa 4680
 |||||||
 Db 4621 GCAACATCGACACAGTGTGCGCTTCAGCGACAGCGGCTGCGCAAGACGACACCA 4680
 |||||||
 QY 4681 ccgaggtgctgcgactctcatccgggagcaacgctccgttgagggacaaagaggtcg 4740
 |||||||
 Db 4681 CCGAGGTGCGGCGCACTTCAATCCGGCAACGTCCTCGTAGGCAACAAAGGTCGAGG 4740
 |||||||
 QY 4741 ccgttcgagcagcagctgtctgctgcgacacatcagctgattgttgacaaacacac 4800
 |||||||
 Db 4741 CCGTTCGACGACACAGGTCGCTGCTGCGCACCATCTACAGCTATGCTGACACCAACAGC 4800
 |||||||
 QY 4801 tgggtctgcaagacctggtgacggttgacactcgacgctaggcgaagagatcaatgca 4860
 |||||||

Db 4801 TGGTGTCTGCAAGACCTGTGTGACGGGTGACACTCCGACCTGAGGAGATCAATGACGA 4860
 |||||||
 QY 4861 tgcgcgtctacgaactaagccgcgctctccgctgtgtgctgagcttctctctgctactg 4920
 |||||||
 Db 4861 TCGCGCTTACGCACTACCGCGCTCTCCGCTGTGCTAGGCTTCTCTGCTCGCTGACTGC 4920
 |||||||
 QY 4921 tcttcgactgtctctgtgagcaagaagcactgaacccgagctgcaagcactctcatga 4980
 |||||||
 Db 4921 TCTTCGCAATTTGCTGTGAGCAAGAGCACTGAACCGGCGACCTGCAACACTCTTCATGA 4980
 |||||||
 QY 4981 cgaagtaagctgtgtgagctccggcggtgaaacccggcgcggtgctggcgcgacacaa 5040
 |||||||
 Db 4981 CGAAGTACGGCTCGGTGAGCTCCGGCTGAACCGGCGGCTGCGCGGCGGACCAACA 5040
 |||||||
 QY 5041 agcaagcgtcgaagggctcccaaccccatgttgaaagagccatccgcgcgcgacttgc 5100
 |||||||
 Db 5041 AGCAGCGCTGAGAGGCTCTCAACCCCATGTGGAAGAGAGCCATCGCGCGCGACTTCG 5100
 |||||||
 QY 5101 acgcatacgtgacgagatggtgcactcgcacactgcatcgacatgagacatgcccgaat 5160
 |||||||
 Db 5101 ACGCATCAGTACGAGCGAGTGGCGACTCCGACTGATCGCATCGAGCATGCGCGCAAT 5160
 |||||||
 QY 5161 tccgcgaagcactacttccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 5220
 |||||||
 Db 5161 TCCGCGAGACTACTTCCCGCGCGCGGACACAGACTCAAGAGCGGCAATCGCTTGACA 5220
 |||||||
 QY 5221 tggcggaagcagcagacaaagcccgctgacacgacatgacacacacttgcgttcaag 5280
 |||||||
 Db 5221 TGGCGAAGCCACGAGCAAAAGCCGAGACACGATGAGCAACAACTTCGGGTTCAGT 5280
 |||||||
 QY 5281 ccaaccgctacctgcacagcgccacccaagtaactcagggatataactgttccagg 5340
 |||||||
 Db 5281 CCACCCTGACTGCGACAGCGCACCCAAATGATCTCCAGGATTAACCTGTCAGAG 5340
 |||||||
 QY 5341 tgcctacgcgcggaagtgcgacacagcggttatacgcggaagaacatagatgaagat 5400
 |||||||
 Db 5341 TGCTTACGCGCGGCAAGTGGCACACCGCTTATCATTCGGGAACATTAGCATGAAGT 5400
 |||||||
 QY 5401 acctatgacatattgtaaatattgatacatatctatttatacaatatatttat 5460
 |||||||
 Db 5401 ACCTATGATACATATTGTAATGTAACATATCTATTATTAACAATATATTATTAT 5460
 |||||||
 QY 5461 atttgctaaaaaataaaaaaataaaactcgag 5498
 |||||||
 Db 5461 ATTTGCTAAAAAATAAAAAAATAAAAAAATCTGAG 5498
 |||||||
 RESULT 2
 AF319973
 LOCUS AF319973 5582 bp mRNA INV 05-DEC-2000
 DEFINITION Manduca sexta cadherin-related protein receptor Br-R1 mRNA,
 complete cds.
 ACCESSION AF319973
 VERSION AF319973.1 GI:11545673
 KEYWORDS
 SOURCE
 ORGANISM
 tobacco hornworm.
 Manduca sexta
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
 Ditrysia; Spingioidea; Spingidae; Spinginae; Manduca.
 REFERENCE
 1 (bases 1 to 5582)
 Dorschner, J.A., Maaty, W.S.A., Griko, N.B., Candas, M. and Bulla, L.A. Jr.
 A cadherin-related Protein Receptor, Br-R1, in the Midgut
 Epithelium of Manduca sexta Mediates Toxicity for Bacillus
 thuringiensis CryIIa Toxins
 JOURNAL
 Unpublished
 2 (bases 1 to 5582)
 Dorschner, J.A., Maaty, W.S.A., Griko, N.B., Candas, M. and Bulla, L.A. Jr.
 Direct Submission
 Submitted (07-NOV-2000) Center for Applied Biology, University of
 Texas at Dallas, Richardson, TX 75083, USA
 FEATURES
 Location/Qualifiers

source 1..5582
/organism="Manduca sexta"
/db_xref="taxon:7130"
CDS 197..5530
/function="receptor for Bacillus thuringiensis CryIA
toxins"
/note="midgut epithelium cadherin-related protein
receptor."
/codon_start=1
/product="cadherin-related protein receptor BT-R1"
/protein_id="AAG37912.1"
/db_xref="GI:11545674"
/translation="MAVDRIAAFLVFIAPAVLAERCGYMTAIPRLPRDNLPLN
FEGTWSORLLPAPERDLCMDAYHVTALNGQVITIMEDEIEITLITLNTGSPV
TPETPLSSSYNLMPIVRDNGEHLITTOQHLEIDMGQMYMNVYDQSLV
AGVSLAIVNDNAPITONFECRVPELGEGLTCTYQVSDARISTEEMTIDIS
VRGDETYIERITPNQMMMLNMTIGVNTSLNVTSLHLEFVALDLPNTHTVM
MVQAVNYSRPRMLEFAVOQFEKSYQNTFVARIDDEITNMEINRYLITNEEDT
FSIEALPGSKGAVFLSPIDRLQREPLTVAAYKDEASTNNVYIYDIN
DORPEPIKEXRLAIMEETPLTNDFKFEGRHDKLQNAQYTVVLESDVPGAFAF
YIAPEVQYRQTFINGTLNHSMLDVEPEQSITRYATPNDNRHRYGALVYHDLI
NMNDQPIFEHNAVOTYTFDEBEGEFPAKAVAHDRIDGYEHTLLNAYVLTIDK
LIGDIRVANDSENFHRESELVQYRAVDLTGEPRHTASQLVLRINDINPTPLRL
PRGSPQVEENVDPGHVITQELRADDPDTADLREINDTSFATKQGHANDPERNC
VEIETIPEIINNRLAIGRVARERIRHNTIDYEEFVSLTVRVDINTYGDYDE
SMLTITIDMNDNPVMEGLEONFVRREMSAGLVYGSVADIDGPELNOYVYTI
FPREDQKDLIMIDFLGOISVNTSGAIDADTPPEPHLYTVASDRCSTEDRDCPP
DPTVETEGNTITHTIDNNKVPQAEETKEPTVYIYENATHLDEVTLLASDLRDE
IYHTVSVINAVNRLNMFNVNREGLVYVDEYEQSGEVLDRGDEPHRIEFL
IDNMEGEGRNQNDNEVLVILDVNDNAEELPPSELSTISNLKOGVRLPEHFI
APDRPEPDNSRVGEILNSTERDIEVPELFVIOQANTGELEMDJKNYGYT
AIIHRAFDHGIPOMSNETEYLIIHPEVYAEPEFINDVIRLARERAYINGLAT
VNGEIEIISATDPDGLHAGVYTFQVYVDEDSORFOVYNGENIGSRILQADPEI
REFRTIATPOGDPGLSDMTFVYVFTQGEPRASSEHNAVLEKRSABESH
QLPADIKNLCEDDCHSITRIIDGSEGHFGLDPEVNRNLFKLKELIRQSSHTL
QVAASNSPDGIPLPASLITVYVREADPPRVRELYRGLYISADISGELRLHA
TQSESAITVAIDYDITVYVREADPPRVRELYRGLYISADISGELRLHA
TATPADGQDRTVTVYVSONRYEFVNTLOOEVRDIEIDTFEFGNMCNID
OVVANDPQVTVALEHSTOMRGHFRDNPVLADEIRDLVIRDSLISQITLARS
VLQDLITNSSPDSAPDSLTYYVLAISAVIGMCLVLLFTITRLANRLRLSM
TKYGSLSGSLNRAGIIAAGTNNKHVESNPLFNEKIKPDDAIDASEGNSDDLIGIED
LPHFGNFMDEVEKANGPYEVANHNHNFNPTPEFPEVNGQFRKI"

BASE COUNT 1478 a 1499 c 1405 g 1200 t
ORIGIN

Query Match 36.7% Score 2017.8; DB 3; Length 5582;
at Local Similarity 64.1% Pred. No. 0;
Matches 3253; Conservative 0; Mismatches 1767; Indels 57; Gaps 12;

QY 500 gggccctgagtttcgacagaatgaagcagcgcccaatgagccttgttattacgcaag 559
Db 547 GGTGATGCCGGTCATGAGAGATTGACACGGGGGTGGCATCTCATCATCCGCAAG 606
QY 560 acaagactacgagacagacacacatgagactatggtctcaatccaatgagagtgta 619
Db 607 ACAGCATTTACAGAGTTGCCGGCATGACAGCATGACATGTTTCATGTGCGGTGACGCCA 666
QY 620 atcaagacgctgactggtggtgcgttgagatagtaacatcagacgacacccgactct 679
Db 667 GTCGCTGGGCGAGCGGTCTCTCTGCTATCGTCAACATAGATGACAAAGCCCATCAT 726
QY 680 gaaagtgtgcaagcgcctgtgtaattccagaacatgagcgaggttagactgaccgactg 739
Db 727 ACAAACTTCCAGACCTTGCCGGGTCTTCACTGAGGCGCCAGGAGGTTACAGAAATGAC 786
QY 740 gtaccaagtgtcagacccgagcgtgtaaacacgacccgcttcatagcttcggtcga 799
Db 787 ATACCAAGTATGAGACGGGACGGAGCGATGACACAGAGTTTCATGAGCTTCAGATGCA 846
QY 800 cagcagcagggctgcagatgaaagcaatcttcacatggttggagaatcagaccacgaga 859
Db 847 CAGCGTTCC---TGGCGACGAGAGACCTTCTACATCGAACGAGCAATATCCCAACCA 903
QY 860 ctggttcaatatgaagatgactggtggtggaatcgccttgaaacttgagacaactca 919
Db 904 ATGAGATGTCCTTAATATGACATAGCGCTTAATACCTGCTCAACTTGTCAACCACTGC 963
QY 920 gctcatatatattagcgtcacagccttctactgcgtcaccaacacacagcgtcacat 979
Db 964 GGTGATATATTTACAGGTGACAGCCCTGACCGTCCCGAACCACCAACGGTACTTAT 1023
QY 980 gatggtgcgaagtggagacgtagagctcgcgcctccgctggtggtgagatcttctcagt 1039
Db 1024 GATGGTGCAGTGGGGAATGTGAACAGCCGCTGCGCGCTGAGATCTTGCTGCT 1083
QY 1040 ggaagagtttgcgagagacgaactcaagagcttcccccgcgcgagatgagggagacc 1099
Db 1084 CCACAGTTTGAAGAGAAATCTTACCAAACTTCCACAGTAGAGGGCATGACGGAGACAC 1143
QY 1100 ggaatcaataagggcatcaactataccctcacaaggatgtagcgtcagacatctcttc 1159
Db 1144 TGAGATCATATATGCTATACATACAGGCTGATCAACATGAGGAAGNACATTTCTGAG 1203
QY 1160 cctggaagtgatgty-----aagacgagactatctcgcgtgactgtagcagccg 1210
Db 1204 CATTTAGGCGCTGCTGCTGTAAGAAAGCGGGGCTGATTCCTGCTGCGCAATTGACCG 1263
QY 1211 cgacaagcttgaagaagagcttccaactcacaactcgtgtgttcaaatcaactgagcgc 1270
Db 1264 CGACACACTGTCACAGAGAGGTTTCCACTTACGATCGTCTTCAAAATATGATGAGGA 1323
QY 1271 tagcttgcacaagagggccacatttcacatcgtcaacgagcgtcaatgatacagacc 1330
Db 1324 GGCCCTTTCACATACAAACAAAGTGTGATCATTTGTGACAGACATCAAGCAAAAGAC 1383
QY 1331 cgaagcgcgtgcaataagaatagatgatatcatctggaagaaactccaatgactctaa 1390
Db 1384 TGAACCTTATACACAGAAATATCGACTGGCAATCATGAGGAGAGCGCCCTGACCTCA 1443
QY 1391 ctcaatgaagaatttgatcttcacatgacatgagatattggttgaaaaacgctcaataacgt 1450
Db 1444 CTTGATTAAGAAATTCGATTTTCATGATTAAGATTTAGGTTTCAAAAACGCTCATACAGGT 1503
QY 1451 ggaacttgaagagcgtgtcccgccagggcggtccgcatctcactcagcgcggggag 1510
Db 1504 GCGTTTAAAGAGCGTGGACCTCCAGGGGCTGCTGAAGCATTTCTACATAGCGCCCTGAAGT 1563
QY 1511 cggctacagagagcagacacttcaatcattggtgacacataaacaacacatgcttgattaga 1570
Db 1564 CGGTACACGAGACAGACCTTCATCATGAGGACCTCAATCACTCCATGCGGATTAAGCA 1623

Oy	1571	agaagtcatttttcgaaacatcatatlaagtcgaagcagltgacatlgaaacacgctag	1630
Oy	1571	agaagtcatttttcgaaacatcatatlaagtcgaagcagltgacatlgaaacacgctag	1630
Db	1624	AGGCCAGAGTTTTCACAGATTATACGGTTTCGGGTGGTACCGACCGCAACAACGACACGAG	1683
Oy	1631	ccacgttgagcagagcgtgtgttgcgtgagaccgtgaaccgtgaaccgtgaacgcgaacgtaccat	1690
Db	1684	GCACGTGCGGCGTGGCTTGGTTTCACATTGACCTCATCAATTGGAAACGATGAGCGCCGAT	1743
Oy	1691	cttcgagagagcagctactccgctgcgttcttlaagagacgctcgcgcggtctccggt	1750
Db	1744	CTTTCGAACACGCCGTGCAGACCGTCACTTCCTTCAGCAGACTGAAAGCGAGGGGTCTTCGT	1803
Oy	1751	ggccacggtgctgcgtccctgcgaagagacatgcgaacgtagtagtgcattatgtatggg	1810
Db	1804	CGCCAAAGCGCGTTGCACACAGACAGAACATCGGGGATGTGTCGAGCATTCATTATTGGG	1863
Oy	1811	caacgcgttgactactcgttctcagaatgaatcaacgcggagagatccttcgtgcgatgga	1870
Db	1864	TAAACGCTTTAACTTCTCTGACCATGCAAAATCTCACCGGCGACATCGCGCTCTACGCTAA	1923
Oy	1871	cgatgcctctgcactaccacgcgaagacaactatcttctcaagltgcgcgttcgacgtac	1930
Db	1924	CGACTCTTCCTCACTACCACTGAGAAAGTAAATTATTGGCAGGTGCGAGCTACGACAC	1983
Oy	1931	tttggtgcagcgcgcccaacaacaacagltgacccaacgcgtgtgtatagaactgcgagatgt	1990
Db	1984	GCTGGGGGGA--ACCTCTCCACACGGCGCACGTCACACGTGGTCATACACATTAATGACAT	2040
Oy	1991	caaaaaacacccccacacctaagctgtgcgcgcctgcgcgtcccaagcgtgcgagagaaagt	2050
Db	2041	CAACAAACGCCACCCACACTTACGGCTGCTCTGAGGACATCCCAAGTGGAGAGAAAGCT	2100
Oy	2051	tcgccgaagatgcagagatalccgggggaaatcaactgcctacacgcacccggaacacagccta	2110
Db	2101	GCCTGATGGGCACAGTCATACCCAGGAGTTAGCGGCACACGCCACCCGACACAGCGCGA	2160
Oy	2111	ccgtgtgtctcgagatcgcactggaactcaactcgtggccaccaacgaagcaggtcagagaccaa	2170
Db	2161	TCTCCGCTTGGAGATTAACCTGAGGACACCTCTTCCGCCAACGACAGGCGCCCGACGCTAA	2220
Oy	2171	ccctacacgaatacgtcgggtgatagtattgcgaacgataatccccacacggggcaacgcg	2230
Db	2221	CCCCGACGAGTTTAAAGAAATGGCGTGGAAATCGAGACACATCTTCCCGAGATTAAACACCG	2280
Oy	2231	gggtctccgcacatcggcgcgccctcgtgtgtgcgaagatlcggcggaacacgtccacatcgcact	2290
Db	2281	GGGACGTGGCTATGGCGCGCGGTGTACCGCGCGCAATACAGACACACGTGACATATGACTA	2340
Oy	2291	cgaggaatcgcgaatgcttactcctcaacgcgtccgcgtgtgaggaactcaaacatgtcatcgg	2350
Db	2341	CGAGAGATTTCAGAGTCTCTCTCTCCACAGTGAAGGGTGGCTGACCTTAAACCGCTTACGG	2400
Oy	2351	agatgcatacgaatgagagcgacgtctcaacgtccaaataacgcgaatgcgaacgcgaacgcgc	2410
Db	2401	AGAGAGATTCAGAGATTCGATGCTCAATTAACATTAATCGATATGAACGACAAACGCGCC	2460
Oy	2411	gactctcgcaacgcgcacgctgtcgcacgcgaacgcgaatgcgtgtgcgcgcgcgtgcgcgcg	2470
Db	2461	GGTGTGGGTGGAAGGACCTGTGGAGCAAACTTCGAGTCCGCGAGATGTGGGGGGGG	2520
Oy	2471	caagctcatcgctcgtgtgtcgcacgcgaatcgcgaacgcgcgcgtctcaaaccaagtgcg	2530
Db	2521	GCTGTGGTGGGTGCGTGGCGCGGACGACATCGACGAGCCGCTCTCAACCAAGTTCGG	2580
Oy	2531	ctaacatatacaactcagaacaacactcccgaggtatagtgaagatgtgacttcacaac	2590
Db	2581	ATAACACATTTCCCTCTGTAAACACAGATTAAGGACCTGATTAAGATCGACTTCCTCAC	2640
Oy	2591	tggtcacaattgaggtgtgacgtgcgaacgcgcgtatgcatagcgaacgcgaacgcgtgcgtctta	2650
Db	2641	GGGTCAAAATTTCCGTGTACACAAAGCGCGCATCGACGCGGATACTCTCTCACCGTTTCCA	2700
Oy	2651	ctgttactaacgcgtatcgtcgtgc	2710

Db	2701	CCCTACTATACAGTGGTGGCTTGTACCGGATGCTCGACAGAAATCCTGGCAATTTGCC	2760
Qy	2711	ttccagattccaaactacttcgaagtctccagcgatatcgaaatagaatcatcatcgaca	2770
Db	2761	CCCTGACCCGACTTATTGGGAAACCGGAAGAAATATCATCAATCCACATCCAGCACGAA	2820
Qy	2771	caacaagatgctcgagccgctcactctgaagaattcaacchogaagtggtgcgtcgtggagaa	2830
Db	2821	CACCAAGGTCCCGCAGCGGAAACGCACTAAGTTTCGTATACCCTGCTGTATATTTACGAA	2880
Qy	2831	tgccagagcgcgagcgaggtggctccagctgtaacctcccgacgctgcacagagcgagtt	2890
Db	2881	CGCAACCACTTATAGCAGAGTGGTCACTGTATAGCAATGATCTTTGACAGAGAGAAAT	2940
Qy	2891	gtacacacggtacgatacacgatagaacttgcgttgaaccccgagcttgcggattctt	2950
Db	2941	ATACCAACAGGTGAGTACGTACTCATCAATTATGCAAGAACCTCCGACTGATGAATCTCT	3000
Qy	2951	cgaggtgaccttggaacactgtgtcgtcttgaagtgcatctaccg-----ggagga	3000
Db	3001	CTCCGTGAACCGAGAGACCGGCTGGTGTACGGCACTGTGAGACCAAGGTAAGTGCGGA	3060
Qy	3002	aaatttgaccgcgatbvggatagagctcaacataactcttgttaaatcatcgata	3061
Db	3061	GGTCTGTGACCGTGATGGTATATACCAACGACCGTATCTTCTTCACCTCATGACAA	3120
Qy	3062	ctctctttcgtatgttgaaggtlagaagtaaacacgagacggaagtgtgaatatattgtcgtct	3121
Db	3121	CTTCATGGGGGAAGGAGAAAGTAAAGAAATGAGACGACAGAAAGTTCTGTATCTT	3180
Qy	3122	atgtgatgtgaagacaagctctccgtgaatgcatgtcgtctgaatgcggtttgaagt	3181
Db	3181	GTTTGAGTGTGAATGACAAATGCTCTCTTAATTTGCCACCGCAGCAACTCTTGTGACTAT	3240
Qy	3182	ttccgaagagcgagttgtcgtgtgtccgtgtgtaacctccgaatactcagcccgagaaagga	3241
Db	3241	ATTCGAAGACCTTAAGCAGGGGCGTCCGTTTGACACCAATATCTTGTGCCCGGACCGGA	3300
Qy	3242	tgaaccagacagcagacaacccgcgtgtcgtgtltagcgaatccctggaactcacaatcagta	3301
Db	3301	CGAGCCCGACACAGACAAATCCAGAGGCGGCTACGAGATCTGAAACCTC--AGCACGA	3357
Qy	3302	ccgagacatcgaggtgcggaatctctcaacatgatctgatttgaataaacaacatcgggga	3361
Db	3358	GCGGGAATCATGAATGCGCGAGCTGTTTGATGATACAGATCCGGAAAGTCCAGGAGA	3417
Qy	3362	acttgaacccgtatgtaacttgaagggtgatbvggtgcattcagaaatcatctatgagcc	3421
Db	3418	GCTGAGACCGCCATGGAACCTCAAGGGAATTTGGGGAGCTACGCTATACATATCGGCG	3477
Qy	3422	cttcgacaacggtactaccgagcagagaggttccaaagaggtacacactgtgtatccgccc	3481
Db	3478	ATTGACACAGGGATTCGCGAAATGTGCCAAGACGAGACATATGAGCTGATCATCTATCC	3537
Qy	3482	ctacaactccacaaccgttgttcgtgtgtcccgacaaccggaactccgtatcgcgtctc	3541
Db	3538	GTTCACTACTACCGCGCTAGTTGCTTCCCTCCGACCAACGATGCCGTATACGACTTGC	3597
Qy	3542	taggagagcgcaacagaagcgcggtcttcggagagcgctgcacaagagtcttcbbagcc	3601
Db	3598	GAGGGAACGACCTGTATCAATGAGATTCTAGCGAAGTGAACGAGAGGAGATTCTTGAGCG	3655
Qy	3602	gattacagccacgagagagggcctccacgcgggcaggttcaagttccacgtccaggg	3661
Db	3658	GATTTCGGCGACTGATCCGAGCGGACTCCACGGGGCGTCTACCTTCCAAAGTGCTAGG	3717
Qy	3662	aaatgagagagcgctltcagtaacttgtataaactgaagtggagcagagaaataagcgg	3721
Db	3718	CGATGAGGAATCAACAGGTACTTTCAAGTAGTTAAGATGGC-----GAGAACCTCGG	3771
Qy	3722	gcagctatattacgcagctttcccgagcaaatcagacacattcaggatacgtccg	3781

Db	3772	CTCGTTGAGGTTTACTGCAACCCGTTCCAGAGAGATCAGGAGATTCCGATATACATTCC	3811
Qy	3782	ggccacggaacggcgccagcagaccggcccgcttggacgcagtcacgttctcgtygt	3841
Db	3832	CGCTACAGACCCAGGGAACGGACCAGACCGCGTGTCCAGGACATGATCGTCAAGTGT	3891
Qy	3842	cttgtaaccaccaaagggcgaccacagtyttcagcgaaaatgcagctactgttcctctt	3901
Db	3892	TTTTGTGGCCACGGCAAGGAACCTAGATTCCGGTCTCAGAACATGCTGTCGTTTAT	3951
Qy	3902	cgaggggtgaagaagcctccgttagagattttgacgtgcgcgaagcgaagaccttaaa	3961
Db	3952	AGAAAAAGAGTCCGGCATTTGGAAGAGTCTACCAACTTCTCTTAGCACAAGACATCAAA	4011
Qy	3962	ccaacctctgcaagaatgactgcgaagatalctactcaagtyttattgaagcgaacaaga	4021
Db	4012	CCATCTCTGTGAAGACGACCTGTACAGCATTTTACATGTATATGATGACACAGCGA	4071
Qy	4022	gggtcttctgtactggaaccaagtcacagcgttcatctcccttgcgaggaattggaccg	4081
Db	4072	AGGTCATTTCCGGCTGGATCCTGTTGGCAACAGTGTTCGAAAGAAAGACTGTATAG	4131
Qy	4082	cgaggtggccagcgtctcaacgcgttcacatcgcgagcgagcgaactccgcgaagccaatg	4141
Db	4132	GGAAACAAGTCCCTCCACACACTGTGCAAGTGGCGGCTAGTAATCTGCCCGATG--GTGG	4188
Qy	4142	gataccctctgagaattccactcctctgttgcagcgttcaatgtgaagagaagcgaaacgcg	4201
Db	4189	CATTCCACTTCTCTCTTCATCCCTTACTGTACATGTTACCGTAGAGGAGCGACACCTCG	4248
Qy	4202	cccaatttcgagcaggaaccttcaacagcgggcatcttcgacgttggaaagcatlygcg	4261
Db	4249	TCCAGTGTGTTGTGAGGAATTTGTACACCCGAGGAGATTCACAGCGGACTCCATCGCAG	4308
Qy	4262	ggaattcttacttgtaagggcgagccaacaaagaagcgacacatcaagtacacataga	4321
Db	4309	AGAGCTGCTCAGATTACATCGACCCACAGTCTGAGGCGTCCGCAATTACTTATGCTATGA	4368
Qy	4322	ccgtgcgagcaatgcaagcttgacagcagacctagaagcgtgcgcgactcgccctgcgct	4381
Db	4369	CTAGAGTATACATGGTATGTGAGACCCGACCGCTGGAGGCGATGAAACAGTGGCGTTTGTACT	4428
Qy	4382	gcatlgcaacacgcgctgtcttctgcataatgcaagccacgcgttccatlygcaagcat	4441
Db	4429	GAAGCGCTCAAAACGAGATGCTGTACGCTTAAATATCCAGCCACGGCGACAGATGACT	4488
Qy	4442	gttcgagttcgaagtcactgcgtcgaatacagcttctgcaatcgacacaaagccgtytga	4501
Db	4489	GTTCAAAATTCGAAATCTCACACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	4548
Qy	4502	agcttacctatctcatcgcgaaaacgcgttgacattcatcttttcgatacgaacttgagac	4561
Db	4549	CGTGTACGTGGTATCTCTCGAGAACCGGCTTACTTCTGTTCTGTCAACACGCTGCACAA	4608
Qy	4562	cgttgcgagaagaacaaattcatalagcggccacgttccagcaacgcggttcaacatgacgtg	4621
Db	4609	GGTGGAGACAAACAGAGACTTTTATCCGCGAACACTTCAACGCTGGGTTCAACATGACTG	4668
Qy	4622	caaatcgaccaggtgtgtgcgttcaagcga----cagcagcggcgtgtgcgaagaagcac	4678
Db	4669	CAACATCTGACCAATGTGTGTCCTCTTACACACCCCGCTCACACCGCGTGGGCGTAGACAG	4728
Qy	4679	caccgagtgctgcgcgaactatccggggaacaagtcaccgcgtgcgaagcgaagaaggtcga	4738
Db	4729	CACCGAATGCGCGGCACCTTCTATACGGGACAAAGTACCCGTAATCTCGCTGATGAGATGA	4788
Qy	4739	ggccgtccgcaagcgaacggtgtgtctgtgcacacatccagctgatatgtgaaacaaacag	4798
Db	4789	ACAATATCCGTAAGTACTACTTCTCTCTGTAGCTTCGATACAAACAAACGCTGGCGGCGATC	4848
Qy	4799	ctgtgtgtcgaagacctgtgtgacggtgtgaacatccgaagcgtatggcgaaggtacaatga	4858
Db	4849	GCTGTGTGTGAGACTGTTGTGACCAATCTCACGCGGACCTCCGCGCGCTGACTGAGGC-	4907

[illegible]

QY	1482	gqccgcgattctatcatcgcgcgcgggagcgcgttaccagagcgaccttcatlygac	1541
Db	1450	gcgaagagccttttttcataatgacaccgcgaagtttggtgtaccacgcagaccccttcattcgtggg	1509
QY	1542	accataaacccaccacgctgcgcgtttacgaagatgctcatcttttcagaacatcatatga	1601
Db	1510	ACAGCTAACCTTAATAATGCTCGATTATGAAGTTCAGAAATTTCCAGGATTCAGGTCAGGTTGGCG	1568
QY	1602	gtcaaaagcgtgagatcagtaaaccaacgcgttaagcaacgctgtggcagcgttgytgacltgac	1658
Db	1570	GTAATATGGAACAGACATGAGCAATGGAAGAACACAGCTGGGGTTGCCGATGTGTACATCAAT	1629
QY	1662	ctgatacaacttggaaacgagaaacttcccatcttcagagagagcagctactcgcgtctt	1721
Db	1630	CTTATCAACTCGAATGATGTAAGAGCCGCTATTTTGAACATACGCTACAAAACGTTTCTTTC	1689
QY	1722	aagagagacgtcgcgcgccgcgtctcccggttgcgcaggtgtcgtcccttcagcagaaatc	1781
Db	1650	AAGGAACCTGAAAGGGAAGGAGATTTTTCGTGGCTTAATGTAAGAGCCACAGCAGAACTT	1749
QY	1782	gaacagcgtagagagcactatgatgtgggcaaacgcgttgactactcgttcaatagaatga	1841
Db	1750	GATGACAGAGTTTGAACACACATTAATGGAAGAACCCAAATTAATTCTCTCGATATATAG	1809
QY	1842	tcaacggtgagagatcttcgtgagatcagatgctccttcgactaccacgaaagaaact	1901
Db	1810	GACACAGGGGACATTTCAATGTGTCAGCAGAGAGACGCTTCTTGACTATACCGAATAAGCCAA	1869
QY	1902	ctatttgttcaggtgtgcgcgttgacagatactttgggcgcagcgccecaacaacagtgac	1961
Db	1870	CTTTTGTTCAGGTCACAGCTGACAGACCCCTGGGGGA--ACCATTTTCACACGGCGACT	1926
QY	1962	accagcgtgtgtagaagacgtggaggtgtgtaaaaaaccccccacacctaagctctgccc	2021
Db	1927	TCTATGCTGCTGATTCATCTGGAAGATATTAACACACCCCTCTACTCTGTGAGATTGGCG	1986
QY	2022	cgttgcactccaaagcgttcgagagaaacglttcccgaagatagagataaccgggaaatc	2081
Db	1987	CGGGCAGTCTTATGTGTCAGAGGAATGTGCCGAGGGCTACATCATCAAGTGAAGTC	2046
QY	2082	actgtctacgaccccgagacccagcgcctaacctgtgtgttcagatacgcacttggactcaac	2141
Db	2047	CGACCCACCGATCTCTGACACTACGCGCTGAACCTGCTTGAATGATGACGTGGACCCCTTCC	2106
QY	2142	tgggcacacgaagcgggcgcagagagaccacaacccactgaatacgtcgggtgtatagtatc	2201
Db	2107	TACGCATACCAACAAAGGACAGGAGCTTAATCTATTGAATTTCCAAATTTGTCTCAGATTC	2166
QY	2202	gaacagatataccccacacggagggcaaacgcgggtgtccgcacatcgggcgcctcgttggaa	2261
Db	2167	GAACATATATATCTCGATGATCAACAAACCGTGGCTGGGCTATAGAGACGCTCGTCTCAAA	2226
QY	2262	gagatccggagacaacgtcaaccatcgacttcgaggaattcagatgcttacttccatccgct	2321
Db	2227	AAAATCAGGGGAACGTCACAATTTGACTACGAAGATTTGCAATTTGTATCTGACGGTC	2286
QY	2322	cgcgttggaggaacctcaacactgtcatcgagatagatcagatcagagcgaggttcaagtc	2381
Db	2287	AGAGTGGGGACCTTAACACTGTTATTCGGGGATGATTAATGATGAATCAACCTTCCACATC	2346
QY	2382	acaaataatcgacatgaaacgacaacgcgcgactcttcgcgaacgagacgctgacgcagcg	2441
Db	2347	ACAAATCATCGATATGAAACGATATATCTCGGATATAGGTGCCGGGAACACTGGAGAGTGC	2406
QY	2442	atgcgcgtgtgcgagctcgtgcgcgcagcgggcaagctcaatcgtcctcgtgtcgtccacgcgac	2501
Db	2407	CTGGGTGTTGGGAAATGTAGTACAGCCGGGTGTCGTGATTTGGGACATCTACGGCCACTGAC	2466
QY	2502	atgcagcgcgcgcgtcaaaccaacatgctcgtcatcatcaacccaacagaaacaacatccc	2561
Db	2467	ATTCAGCGGACCACTTCAATATCAATGAGTCATACCTATATAAGGCAACGAAGGCACACCA	2526
QY	2562	gagggatagtgaaagatgtgcttcaacaactgttcaaatgtgaatgtgacgtgaacagagcg	2621

Db	2527	GGAATCTATTGAGATCGATCTTTCACAGTCGGTCAGATCAAGGTGAAAACAGCGCGCA	2586
Qy	2622	atcgaatgaagacgaacccctggcgtcttactctgttaacacccgtcaatcgttagaagag	2681
Db	2587	ATAGATGCTGATGTTCCACGAGATACACCTTACTACTCTGTGGCCACTGATCGG	2646
Qy	2682	tgctccctcggaaaacgcgaagaaatgctccctccagatctccaactacttcgaatctccagc	2741
Db	2647	TGCTACGCTGAAGAACCAAGACGATGTGCCCTGACAGACCCGACTTATTTGGGAAATCCCGC	2708
Qy	2742	gatalcgaaataagaatcaatcgcacaacaacaagaatgctcgtgaagcgtctcaatcgtagaag	2801
Db	2707	CAGAGTCGTGATCAAAATCATGATACCCACACAAAATACCTCAACCTTCAACACACCA	2766
Qy	2802	ttcaacacgaagcgtgtaagtcctcgtggagaatgccaagcagcggcgagaggttggtccaagtc	2865
Db	2767	TTTTAAGGGGTCGTGATATTATTCAGAAAGTCGTGTGAGATGACGATTTTCAAGT	2828
Qy	2862	tactccacgcacgtcgacaagaagaaatgacacacggtacgaatatacgaatgaactt	2921
Db	2827	ATTGTTACCGATTTTGGATAGAGATGATTTACACACAAATTCGTTACCAATTCACATC	2888
Qy	2922	gcggtgaacccccgcacgtcggaattcttcgaagtggaacctggaacatgctgcgtctgag	2981
Db	2887	GCCGTGAATCCTCGTCTGCGGAACTTCTTGCTGTGATCCGATACGGAGAGATGATC	2946
Qy	2982	gtgcattccccggggagacgaaaaatgagaccgcgaatggggatagaactcatatctac	3044
Db	2947	GTCATCTACACACCGACGAGATCTCTGACCGGGAGCGGACACACCTTCACACAGATC	3006
Qy	3042	tttgaatttcacgtactacttcttctcgtatggtgaacgttagagaacacgaagaa	3104
Db	3007	TTCTTTAACTCCTCATCGATTAACCTTCTTCACACAGAGACGGCACACGAACCAAAATGAT	3066
Qy	3102	gttgaataattgctcgtctctatctatgtaatgltgaacgaacacgctctcgtagaatgctatgct	3161
Db	3067	GCCGAAAGTACTGGTGTACTCTTGATGTACAGACAAATGCTCCAGAGCTTCAGAACCC	3128
Qy	3162	gatgaacatcggtttagatglttccgaagagcaatgctggtgctgctgtactcccagaa	3221
Db	3127	GATGAACGTGATGTCGTGTCTGAAAAGTTTAACTAAAGGTATCAGCGTCTCCACGACAC	3186
Qy	3222	atctcagcaacgcgacagggatgtaacacgaacgaagaactcggctcgtttagaagaatc	3281
Db	3187	ATATACGCACTGTATPAAGACGAACTGACACGCGCAACTCTCGGGTCCGATACCCATC	3248
Qy	3282	cttgaacctcgaacgacacgcgacgaatcgaagtggtcgcgatctctccacatgactcgc	3341
Db	3247	ATCAGCCTCAGATCGCCCAAGGGAATGGAATGCCAGAGCTTTTCAACATGATACAG	3306
Qy	3342	attgaanaacaaactvgggaacttgagaccgtatagacttgaagggatattggggcaat	3401
Db	3307	ATACAAAACGTCACAGAGAACTGGAACCGCATGACATTAGAGGCTATTGGGGAACT	3366
Qy	3402	tacgaataattcatctggagccttcgacaacgcgtacccgcgacagaggtccaaagaagc	3461
Db	3367	TATGCTATTCTATTAAAGCTTACGACACACGCGATCCCGACAGATGTCCAAATAAACG	3428
Qy	3462	tacaacctgycalcgcgcccctcaaatctccacacacccctgttctcgtgttccgaaacc	3521
Db	3427	TACGAGCTCGTGAATTCGTCTTCAACTTCCACAGCTCTGTTTTGTGTTTCTTAAACAC	3488
Qy	3522	gactccgtcatctggcgtcttctagggagcggcaacgaagcggcgtctctgvcgaagct	3581
Db	3487	GGAGCTACTCTTCACACGGCGCAGGAGAGGGCTGTGTCAAGGGTTTCTGGCAACATGT	3546
Qy	3582	gccaaagatttcctggggacgatctacgcacacacgaagagagcgtccacgcgggcgc	3641
Db	3547	GATGCTGAATTCCTTAAACCGCATCTGCTTACCGATGAAGAGCGGCTACATGCTTGACAA	3606
Qy	3642	gtcaagttccacgycgcagggaaaatgagagggcgttctagttacttgatataactgaagt	3701


```

Db 3607 GTTGCCCTTGAAGTGGTGGTATGATGAGCCGCTGATTAATTTCCACAT-----TGTG 3660
Qy 3702 gtagagagaagaataagcggcagcttatattacgcgaacttctccaggaagaatcaga 3761
Db 3661 AACGACGGGGAACAGTGGCACTGTGATGCTGAACACACTCTTCCCGACGACATTAAGA 3720
Qy 3762 caattcaagatcacgctccggccacgagcggcgacggagccggccggcttltgacc 3821
Db 3721 GAATTTGAGGTGACGATACGCTACTGACGGCGGAGCAAGAACAGACCGCTCCACG 3780
Qy 3822 gacgtcaagcttctgtgtctctgttaccacacagggcgagccagtgctcaggaat 3881
Db 3781 GACCTCACCTTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3840
Qy 3882 gaagctactgtcgccttctcgaaggtgaagaagcctcgtgagaggtttagctgcgc 3941
Db 3841 ACTCACACTGTTGCTTCTTCTTGAAGAAAGAGGGGCTTTTGGAGCTATGAGCTGGCG 3900
Qy 3942 caagcagaagaccttaaaaaaacacactctgcaagatgactgcgaagatatctactaag 4001
Db 3901 CGTGTGAGAGATCGGAAAACCATCTTTGCACTGATGATGTCACAAATATCTACTACAG 3960
Qy 4002 ttatttgagcgaacaacagaggtcttctgtacttgagaccagtcgaagcgtcactcc 4061
Db 3961 ATTATTGATGAATTAATATGACGGCCACTTGGACTGACGAACATGACGAACGTTCTATTTC 4020
Qy 4062 ctgtcgaagagctgagccgaggtgagccagcgtcttaccgctgacacgtcggcgagc 4121
Db 4021 CTGCTTAAGGATTTGGATGGAGTGCACAGCAACGTTATACATGACGATCGCTGGCAGC 4080
Qy 4122 aactgcgcgcagcagcacttggatccctctgcagacttccatccgtgtgtcagcgtcaat 4181
Db 4081 AACTCAC-----CGACAGCGGAATGCTCTAACCTTACATACATACATACATACATACAT 4134
Qy 4182 gtaagaagaagcgaacccggcccaatttctcagcagagaccttaccagcgggacttgcg 4241
Db 4135 GTGAGAGAGGAGATCCCGACGCTTCTGTCAGAGATTTGTGACGGCTGGCATTTTCC 4194
Qy 4242 acgttggagcagcttggccgggaattgtctactgtcagcggcgagccacacagaagcgac 4301
Db 4195 ACTTGCGATTCATACACAGAGACACTGTTGATTTTACAGGCGACACATTCAGAAAATGCA 4254
Qy 4302 accatcacgtacacatagaccgtgagcagctgagcagcagccttagaagccgtg 4361
Db 4255 CCGATATATCTACACGATAGACTGTCAGAGATGATGATGATGATGATGATGATGATGATG 4314
Qy 4362 cgcgactgcgctcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 4421
Db 4315 AGGAGACAGCCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCA 4374
Qy 4422 accgttccatcagcagcgttctgagctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 4481
Db 4375 ACTGCGTTCATACAGGAGATGTTGCAATTCAGATGCTGCTACGATCAGCTGCTGCTAC 4434
Qy 4482 atcgcacacacccgtgtgtaagctcaccatcctcgtcgtcgtcgtcgtcgtcgtcgtcgt 4541
Db 4435 ACTGATAGAGCAAAATGTCAAGATCTACTGATCTCCACAGAGACCGCTTCTTCTTCTG 4494
Qy 4542 ttgcataacaaacttgagacgttgagcagaagaatattcatagcggcagcgttcagc 4601
Db 4495 TTTGTTAACTTTAGAACATCGAGCAGAACTGACTGATACACAAAGCTTTTCC 4554
Qy 4602 accggttttaacatgacgtgcaacatcgacacaggtgtgtcgttgcgttcgagcagcgagc 4661
Db 4555 GCGGCTTGAATGACGTTAACTGACCAAGTGTGCTGCGCCACTACGCGACGCGG 4614
Qy 4662 gtgagcagaacacacacagcaggtgagcagcagcagcagcagcagcagcagcagcagc 4721
Db 4615 GTCATCATTAAGGATCACCAGAGTGAAGGAGGAGCACTTATCCGCTATATGTGCCGCTGA 4674
Qy 4722 cagcagaagaagcgtcagcagcgtcgcagcagcagcagcagcagcagcagcagcagcagc 4781
Db 4675 CCGCGCAGAGATGAGACTCTCCGAGGGGACATGTGCTGCTGACTGCCATCCAGTCC 4734

```

```

Qy 4782 atgctgagcaccacagcctgtgtcgtcgaagacgtgtgacgggtgacactccagccta 4841
Db 4735 ACGCTGGCAGCGCGCTGTAGTGTGCGGACGACTTTCACGAGACACAGTCCGGCGGCC 4794
Qy 4842 ggcgagagatcaatgacagatgcgcgtctcagcactagcagcgcgtcgcgtgtgtagc 4901
Db 4795 GACGCGGGCTC---GGCCCGCGGTGTGTACGCGCTGCGCTCTTCCACCCGCTGGCT 4851
Qy 4902 ttccctgcgtcgttactgtcttctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 4961
Db 4852 GCGCTGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4911
Qy 4962 ctgcaagcactctcactgacgaagcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 5021
Db 4912 TTGGAAGCCCTGACGGTGAAGATGACGATGCGTGGATTCAGAGTTGAACCGCGTGGC 4971
Qy 5022 ct---ggcgccgggacacacagcagcgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 5078
Db 4972 ATGCGCGCGCGCGGACCAATAAATGCGTGAAGGCTCCAAATCGATCTGGAACGAG 5031
Qy 5079 ggcacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 5138
Db 5032 ACCATTAAAGCCCGGACCTTCGATTTATGAGCGATGCAATCAACGACCTGTGATCTAATC 5091
Qy 5139 ggcacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 5194
Db 5092 GGGATTGAAGATGTGCGACACTTCGCGGAAATTAATTAATTTCTCTGTCGTCGCA 5147

RESULT 4
AB041508 5172 bp mRNA INV 11-JUL-2000
LOCUS Bombyx mori btr175a mRNA for cadherin-like membrane protein,
DEFINITION partial cds.
ACCESSION AB041508 GI:9049394
VERSION AB041508.1
KEYWORDS Bombyx mori larva midgut cDNA to mRNA.
SOURCE Bombyx mori
ORGANISM Bombyx mori
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
Ditrysia; Bombycoidea; Bombycidae; Bombyx.
1 (sites)
Ikawa,S., Tsuda,Y., Fukuda,T., Sugimoto,K. and Himeno,M.
cDNA cloning of the CryIaA receptor variants from Bombyx mori and
their expression in mammalian cells
JOURNAL Unpublished (2000)
REFERENCE 2 (bases 1 to 5172)
AUTHORS Ikawa,S., Fukuda,T., Himeno,M. and Sugimoto,K.
TITLE Direct Submission.
JOURNAL Submitted (07-APR-2000) to the DBJ/EMBL/GenBank databases. Takashi
Fukuda, Osaka Prefecture University, Department of Applied
Biochemistry, 1-1 Gakuen-cho, Sakai, Osaka 599-8531, Japan
(E-mail:tfukuda@biochem.osakaifu-u.ac.jp, Tel:+81-722-54-9464,
Fax:+81-722-54-9463)
FEATURES
source
1..5172
location/Qualifiers
1..5172
/organism="Bombyx mori"
/db_xref="taxon:7091"
/dev_stage="larva"
/tissue_type="midgut"
1..29
/feature="PCR primer site"
25..5172
/feature="btr175a"
25..5172
/feature="btr175a"
/function="receptor for B. thuringiensis CryIaA toxin"
/product="cadherin-like membrane protein"
/protein_id="BAA99404.1"
/db_xref="GI:9049395"

```


[illegible]

```

primer_bind      5155. .5172
                  /gene="btr175a"
BASE COUNT      1483 a 1292 c 1225 g 1172 t
ORIGIN

```

Query Match	35.98;	Score 1975.2;	DB 3;	Length 5172;
Best Local Similarity	63.28;	Pred. No. 0;		
Matches 3207; Conservative	0;	Mismatches 1823;	Indels 48;	Gaps 9;

QY	144	aacaaagatcatatgcaaaatgggggtttggaagttctctccaaagatgctactggtctct	203
	7	AACCAAGGAACATCCAAAGTGGAGTTGACGTTCCGAATTCTACCTACCACTTCTAATC	66
QY	204	ttaagcctctgcccagcatagcaacaaacagatgtttcgacattatgcgaatfaccaaagc	263
Db	67	TATGCAGAAATGTTATTTGGCTCAAGAAAGATGGCGCTTCAATGGTTGCCATACCCAGCA	126
QY	264	gagatccggaactgcccgcctatgtatlaacgaagaaatcatfaggaatgacagcctcta	323
Db	127	CCAAGCGCGGATGTGCTCTGAACTGTGACTTTGAAGGTCAACCTGAGTCAAGACCACTA	186
QY	324	ataccggcccgccagccggaaggaagtatgtatgtggaact-----tctaccggat	374
Db	187	ATTACCGCGCGCAGACCGCTGGAAGATGTCGTGATGGACGCGTACCCACGCGATGACACTACC	246
QY	375	caaatgattcaggtcatatcatatggaagaaatcgaaagaaagtcatcatcttggaag	434
Db	247	TACGGAATCTCAGATTATTTTACATGGAAGAAAGAGATAGAAAGAGAAAGTACCCATTCGCTAAA	306
QY	435	cttaactacaaggttcaacaacgccc-----gltgctgtcgaattatgtcaggccagccc	488
Db	307	CTCAATTACCGCGGCTCCAAATGTACCATCATTTGAGCCACAGATTTCGTGTGTTCTTC	366
QY	489	agagcccaagtctgggccttgagtttcgacagaatggaagcagaagccaaatggaagcctgtt	548
Db	367	AATCTACTCTGCGCCGTTATTCGACGAGATACCAAGATAGTAAATGAGAAATGGCACTTATT	426

QY	549	atlaagcaaaagacaagactaagagaaacgaaccaatlgaaagctatgltcttaatcca	608
Db	427	ATTACTCGAAGACAAAGATTACGAACCTCGGGAATGACAGAGATATTTTCAACATCAG	486
QY	609	gtgagaggttgataccaagacgaactctgtgtgagcgttgtagatgaatcaatcagacgaat	668
Db	487	ATGATGTGGGGAGACATTGTGTGCGGGGTGTTTCACTGCTGATTGTCAATATTGACGCACAC	546
QY	669	ccgcccacccctgcgaagtgtgtaacgcctgtcgtaattccaagaacatgycagcgttaacgt	728
Db	547	GCACCTAATATTCAAGCACTTGAAACCGGTGTCAAGTTGACGAAGTAAAGGGAGCTGCTGTG	606
QY	729	accgaactcgtgtaccaaagtgtcagaccgcgaagtguaaatacgaaccgcgtctatgacy	788
Db	607	ACTGTGCTCGATATCGTGTAAACCGAATCCGATGTGGCGCATCGACACCCAGTTTCAACGAG	666
QY	789	ttccgtgtgcacgaacgaaggtcgtacagatgaaagcaactctctacagtgttgtagataac	848
Db	667	TTTAGATTCGACACCGATCG---AGGAGACGACACAGATCTTCTACATTCAGAGGGCGAAC	723
QY	849	gaacccacgcgaactgtgttcaataaagatgactgtgtgggataccaatcccttgaaattc	908
Db	724	ATTCCGGCGGAATGGATCCGATGCCGATGACATGACCGTGGATCAATGAAACGGTTGAATTC	783
QY	909	gagacaactcagctctcatataattagcgtcaacagctcttgaactcgtctaccgaacaacac	968
Db	784	GAGACTATTCACATCACTAATATATTTCTGTGACCGCTTGAGACTCTCTCCCAACACTCAC	843
QY	969	acggttaaccatgagtgtgcgaagtgtggaaagtagtagtcttcggccctcgtcgtgtgtgag	1022
Db	844	ACTGTAACTGCTAATGTATGATCAAGTAAAGTGTGAACATCGCCGCGACAGATGGGTAGAG	903
QY	1029	atcttctcagtcgacgaattgtgacgaagaacaaactcaagactctccctccgcgcgata	1088
Db	904	ATCTTGTGTCACAAACAATTTCGATGAGAAAACGTGCTCAGAGCTTCCCTGTAGAGCTATC	963
QY	1089	gaacgggaacacggaatcaatbagggccaataactataccctcaatcaggatagacgtgac	1148
Db	964	GACGAGATACAGGCATCAACAAACCTATCTCACTACAGACTGGAAACACGACAAAGAAAGAT	1023
QY	1149	gaactctt-----ttccctggaggtgtagtgaagaagggactatcttgcagtgact	1199
Db	1024	ACGTTTTTTCACATTGCGCACCTATGAGAGGGGGTGTGAGCGGGCCCATATTATATCTGAT	1083
QY	1200	gagatcgcacgcgcgaacgcttgaaagagagctttcaacctcaacatcgtctgtacaa	1255
Db	1084	CCAAATAGACAGGGACACATATTCAAAAGAAGAGTGTTCATATTATCCATATGCGCGATATA	1144
QY	1260	tctactcagctagcttctgcacaagagcccacatlltcatcogtloaagaagcgtcaat	1315
Db	1144	TACGACACACGATCTTCAGCTACGGCGCCCAACGTTGTCATATATGTCAACGATATTAAC	1203
QY	1320	gataagagaccgcgacgcgtgcataaagaatacagatattgatataatgtagaanaatcca	1379
Db	1204	GACCAAAAGACCAGAACTTTGTTCAAAGATATCTGTTTAAACATATAGTAAAGAGACAGCT	1263
QY	1380	atgaactaaactcaatgtagaagattgtgatacgaatcgcgaattgtgtgaaagcgt	1433
Db	1264	TTGAGCTGTGAACCTTCGATTCAGGAATTCGATTCGATGATGAGAGATCTGGGTCAAAAACGA	1332
QY	1440	caatlacacagltgaaacttgagaagcgtgttccgcgcgaagggcgcgtlccgcattctacac	1499
Db	1324	CAATATACCGCGTAGAGGCTAGAGAGCGATATTATCCAGACGACGGCGGAACCTTTTATACAT	1383
QY	1500	gcgcgcggagagcggtctacacgaagagcgaactctcatatggtgcacataaaccacacacg	1558
Db	1384	GCACCCCAAGTTGGCTACACGACACGAGACCTTCATCATGTGGAGACAGCTTAACCTATTAATG	1443
QY	1560	ctggaattacgaatgttcaattttcttgaaacatcatcattaagaagtgaagagcgttgacgt	1619
Db	1444	CTCGATTATTAAGTTCCAGAAATTTTCAACGAGATCAGGTTCCGGGTATATAGCAACACACATG	1503
QY	1620	aacaacgctacgaacgltggcgcgcgcgtgtgtgtgtaagctgtaaacctgatacacttgaaagac	1679

1504 GACATGAAGAACGTCGGGTTGCGATGTGACATCAATCTTATCACTGAGATAT 1563
1680 gaattcccatcttcagagagagcagctaccctccgctcgtttaagagacgctggcgc 1739
1564 GAAGAGCTATTTTGAACATAGCGTACAAAACGTTTCTTCAAGAAACAGAAAGGAG 1623
1740 gctcccggtgagcagcgtgctcgccttcagagagacatcagacagatagtagat 1799
1624 GGAATTTTTCGAGGCTAATAGTAAGAGCCACAGACAGACATTTGATGACAGATTACAC 1683
1800 tcatatgtaggacagcgtctgtactcctgttcataagatgaatcaagggagagatcttc 1859
1684 ACATTAATGGGAACGCGCAATTAATATCTCTGATAGTAAGACACAGGGGACATTCAT 1743
1860 gtggagcatgagcagatgcttcgtactacacacagagaaagaaactatattgttcagtgagc 1919
1744 GTACGCGAGAGAGCTTTTGTGACTATCCGACAAACGAACTTTTGTGAGGTACAGA 1803
1920 gctacgaatacttgggagcagcagccacacacacagtgacacacacagctgtgtataga 1979
1804 GCTGACGACACCCCTGGGGGA---ACCATTTACACGCGGACTTCTCAGCTGCTGATTCAT 1860
1980 ctgagagatgtcaacaacacactcctccacacactcgtgcccgttcgaatccaaagctc 2039
1861 CTGGAAGATATTAACACACCCCTCTCTACTCTGAGATTGCCGCGGACGCTATATGTC 1920
2040 gaggaagacgttcccgagagatagagatccggagaaatcagctcagcaccgagc 2099
1921 GAGGGAATGTCCCGAGGCTACATCTCATCAAGTGAATCCGACCCAGATCTTAC 1980
2100 accagcgctcactcgtgtgttcagagatcagctcagcagcagcagcagcagcagcagc 2159
1981 ACTACGCTGAGACGCTGCTTGAAGATGACTGAGCACTTCTTACGCAACCAACAAAGC 2040
2160 agagagacacacactcactgaatcgtcgtgtatagttatcagaaagatatcccaacc 2219
2041 AGGGAGCTAATTCATATGAATTTCAATTTGTCGAGATCGAAACAAATATATCTCGC 2100
2220 gagggagacggggttccgcagctcggcgctcgtgtgtcgaagagatccggagacagtc 2279
2101 ATCAACAACCGCTGGCTGCTATAGGAGCCCTCGTCAAAAAAATCAGGAGAACGTC 2160
2280 accatcagcttcagagaaatcagagatgcttaccctcagcgtccgctgagagacccaac 2339
2161 ACAATTTGACTAGGAAGATTGAGATGTTGATCTGACGGCTCAGAGTGGGAGCTTAAAC 2220
2340 actgtcagcagatgagatcagatcagatgagcagcagctcagatcaacaataatcagatgaac 2399
2221 ACTGTATCCGGGAGATGATGATGAATCAACCTTCACCATCAACATCATCATGATGTAAC 2280
2400 gacacagcgacgcatcttcgagaaagcagcgtcagcagagacagatcgtgtgagagtg 2459
2281 GATTAATCTCCGATATGAGTCCCGGGAACACTGAGCAGTGCCTGCTGCTGGGAAATG 2340
2460 gcgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2519
2341 TCAGACGCGGCTGCTGATTTGGGACTCTGACGGCCATGATGATGAGGACCACTTAC 2400
2520 aaccaaagcgctacacatacacaactcagaacaacactcccgagagattagtagaatt 2579
2401 AATCAAGTGCATACATATGAAAGCAACAGACACACAGAAATCTAATGTGATGATC 2460
2580 gactcacaacagtgatcaaatgtgagtgagcagacagcagcagcagcagcagcagcagcagc 2639
2461 GACTTCTTCACTGGTCAATACAGGTGAAGAACAGCGGCGCATATGATGCTGATTTCCA 2520
2640 tggcgcttactactgtactacacgctcagctagcagcagcagcagcagcagcagcagcagc 2699
2521 CGGAGATACAACTTACTACACTGTGTGGGCACTGATGCTGTAGCTGAAGAACCA 2580
2700 aacgaatgctcctcagatccaaactactcgaagttcagagcgatatacgaataagaatc 2759

2561 GACGATTCCTGACGACCCGACTTATTGGGAACCTCCCGGCCAAAGTCGTATCAAAATC 2640
2760 atcgacacaaacaaagtgctgtagccgctcacttgagaagttcaacacagcagctgac 2819
2641 ATAGATACCAACCAAAATATACCTCAACCTGAAACAGACAAATTAAGCGTGTGAT 2700
2820 gtctggagaatgcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2879
2701 ATTTCGAAGATGCTGTGTGAGATGAGGTGTCAAGTATATGTAAGCATTTGGAT 2760
2880 agagaagcagtggttaccacacagcagcagcagcagcagcagcagcagcagcagcagcagc 2939
2761 AGAGATGATATTTACACACAAATTTGTTCCAAATCAACTACGCGCTCAATCTGCTGTG 2820
2940 cggagattcctcgaagtgagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2999
2821 CCGGACTTCTTGCTGTGTGATCCGATATCGGAGAGGTGATGCTTACTACACACGAC 2880
3000 gaaaaattgagacgcagatgagcagcagcagcagcagcagcagcagcagcagcagcagc 3059
2881 GAGGCTCTGACCGGAGCGGAGACCACTCAACAGAGATCTTCAACCTCATTCGAT 2940
3060 aactccttctcgtatgtagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3119
2941 AACTTCTTCCAGCAAGAGAGACGCAACAGAAACAAATGATGCCAGTACTGCTGCTGA 3000
3120 ctatggatgtgaacagacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3179
3001 CTTCTTGATGTGAACGACAAATGCTTCCAGAGCTTCCAGAACCCGATGMACTGATGTCG 3060
3180 gtccgagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3239
3061 GTGCTGAAGTTTAATAAGGATGACGGCTCCAGCCACACATATATAGCACCTGATAGA 3120
3240 gatgaacagacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3299
3121 GACGAACCTGACACGCAACAACTCGGGTGGATGCGCATATCAGCTCAGATCCG 3180
3300 gaccgagacatcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3359
3181 AACAGGAAATGAGTCCAGAGCTTTTACCATGATATACAGATCAAAACGTCACAGGA 3240
3360 gaacttgagacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3419
3241 GAACHTGAACCGCATGAGCTTGAAGGCTATTTGGGAACTTATGATCATATATTA 3300
3420 gcttcgacacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3479
3301 GCTTACGACCAAGGATCCCGCAACAGATGTCATGAACCTTACGAGCTCGTATTCGT 3360
3480 ccttacaacttcacacacccgtgttcgttccgcgacacccagcagcagcagcagcagcagc 3539
3361 CCTTACAACTTCCAGCTGCTGTTTGTGTTTCTTAAACAGAGGATACCTTGTGACAG 3420
3540 tctagagcgcgcaacagagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3599
3421 GCCAGGAGAGGCTGTGCTCAACGTTTGTGCGCAACAGTGAAGTGAATTCCTTAAC 3480
3600 cagatctacgcacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3659
3481 GCGATCTGCTTACCGATGAAGAGCGCTACATGCTGCAAGTGGCCCTCAATGCTGT 3540
3660 ggaatagagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3719
3541 GGTGATTAAGAGCGCTGATTTTCCACAT-----GTCAAGAGCGGAGAGAACGT 3594
3720 gggcagctataatagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3779
3595 GGCACCTGATGCTGAACAACTTCCCGAGGACATTAAGAAATTTGAGTGAAGATA 3654
3780 cgggcacagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 3839
3655 CGTGCTACTGAGCGGAGACAGAACAGACCGCTTCACAGGACTGACACTTCACTGAGTGTG 3714

[illegible]

OY	4920	cttcctgcattgtctctgttaagacaagacacgtgaacccgacgtcgaagcatctccatg	4979
Db	4786	CTAGTCATCTTTCATCATCAGGAGCTATAAAAACTTAAACCGTCGTTTGGAACCCCTGACGGTT	4845
OY	4980	acgaagtaagctctcgtgtgagactccggcgctgaaccgcgcgcggcgct--'agcgccggagcac	5038
Db	4846	AAGAAGTAGCGGATGGTGGAATTGAGGTTTGAAACCGCGTCGGCATCGCGCGCGCGGAGCAC	4905
OY	5037	aacaagcaagcgcgtctcgagggctccaaccccatgttgaaagagagccatccgcgcgcggac	5096
Db	4906	AATTAACATGCGCGGGAAGGCTCTCAATCCGATCTGGAACGAACCATTTAAACCCCGGAC	4965
OY	5097	ttcgaagccatcaatgaagcgaagtggcgagactccagactatcgcgcatcgagagacatcg	5166
Db	4966	TTGCATTCTATGAGCGATGCGATCAAAACGACTCGATCTAATGGCATTGAAGATGTGCCA	5025
OY	5157	caattccggagcgaactactccgcgcgcgcgcacacaga	5194
Db	5026	CACCTTCGGCGAAATATATTTTCTCTCGAGACGTGCA	5063

RESULT	5
AB041510	AB041510
LOCUS	5172 bp
DEFINITION	Bombyx mori btr175c mRNA for cadherin-like membrane protein,
ACCESSION	AB041510
VERSION	AB041510.1
KEYWORDS	Btr175c.
SOURCE	Bombyx mori larva midgut cDNA to mRNA.
ORGANISM	Bombyx mori

REFERENCE

AUTHORS Ikawa,S., Tsuda,Y., Fukada,T., Sugimoto,K. and Himeño,M.
TITLE CDNA Cloning of the Cyt1A receptor variants from Bombyx mori and their expression in mammalian cells
JOURNAL Unpublished (2000)
REFERENCE 2 (bases 1 to 5172)
AUTHORS Ikawa,S., Fukada,T., Himeño,M. and Sugimoto,K.
TITLE Direct Submission
JOURNAL Submitted (07-APR-2000) to the DDBJ/EMBL/Genbank databases. Takashi

FEATURES

Source

primer_bind

den

22

FYIAPEVGYQROTFIGMTANHKMLDYEVEPEFORIRLRVIATDMDNEEHVGVAYVYINL


```

1684 ACATTAAATGGGAAAGCGAATATATCTCTCGATAGACAAAGACACAGGAGACATTCAT 1743
1860 gtgagcactggaagatgacctctgactacacacgacagaaacactcttggttgaagtgcc 1919
1744 gtcacgacgagacgactcttcttgatcattcaccgaaaggaactcttggttgacgacga 1803
1920 gctgcagatcttggcgagcgccacacacacagctgacccaccgctggtagaa 1979
1804 gctgacgacacccctgggga---ACCATTCACACGGCGACTTCACAGTCTGATCCAT 1860
1980 cttgagagatgtaacaacaactctccacccctacgcttgccccgttcgactccaaagctc 2039
1861 ctgacgagatATAAACAACCCCTCTACTGTGAGATTGCCGGGAGGCTCTATATGTC 1920
2040 gagggagacgctcccggaagatacagagatcccgggaatacctgctacgacccggac 2099
1921 GAAGGAATGTCCTCCGAGGCTACATCAACAGTAGAGTCCGAGCCACCGATCCTGAC 1980
2100 accagcgactctgtgtgagatcgactggaactcaccactcggccacacagagagc 2159
1981 ACTAGGCTGAAGCTGCGCTTCGAAATCGACTGGACCACTTCTACGCAACCAAGAGGC 2040
2160 agagagacacacactactgaaatcagtcggtgtagatgatacgaacgataccacc 2219
2041 AGGGAAGCTAATCTATTGAATTCACAATTGTGCGAGATGGAACAATATATCTCTGG 2100
2220 gagggacacccgggtctccgacatcgggcgctcgtgtgtaagagataccgggaacgctc 2279
2101 ATCAACAACCGGTGGCTCGCATATAGGACGCCCTCGTCGTAACAAAATAAGGAGAGATC 2160
2280 accatcgactcgaggaattcgaatcgatgcttaccctcaccgctcgctgaggaactcacc 2339
2161 ACAATTGACTAGAGAGAGTTCAGAGTGTGTATCTGACGGTACAGTGGGACCTTTAAC 2220
2340 actgctacgagatgactgaaatcagatgagggcgacgcttcacgatacaataatcagatgac 2399
2221 ACTGTTATCGGGGATGATTAATGATGATCAACCTTCACCATCAACATCATGATATGAAAC 2280
2400 gacaaacgacgactcttcggaacgagcagctgacgacgagatgacgctgacgagctg 2459
2281 GATTAATCTCCGATGAGGTCCCGGGAACACTGGAGCAGTGGCTCGTTCGGGAAATG 2340
2460 gcgagcagcgagacgctatcgctcgtctcgacacgacatcggagcccgctcacc 2519
2341 TCAGAGCGCGGCTGCTGATTTGGGACTCTGACGGCACTGACATCGAGCGACCACTTAC 2400
2520 aacaaagtgcgtacacatacaactagaacaacactcccgaggaattagtgagaatc 2579
2401 AATCAAGTCCGATACACTATGAAAGCAACGAAGCACACAGAGAAATCTATGTGATC 2460
2580 gactcacaactgtaaatgtaggtgtagacgaaagcgatcgtatgcagacgaaccc 2639
2461 GACTTCTACACTGGTCAATCAACAGTGAACAAAGCGCGCAATAGATGATGATGTTCCA 2520
2640 tggcgctctactgttactaacacgctcactgctgacgagctgctccctggaacccgc 2699
2521 CGGAATATCAACCTTACTACTGACTGTGGTGCACCTGATCGTACGCTGAACACCA 2580
2700 aacgaatgtctccacgactcgaactcgaactcgaagtccagcgagatataagaaatgac 2759
2581 GACGATTCCTCGAGACGACCAACTTATGGAAGATCCCGGCAAGTGTGATACAAATC 2640
2760 atcgacaacaacaagaatgacctgagcgctcactgagaaatgtaacagacggtgtac 2819
2641 ATAGTATACCAACAAATAATACCTCAACCTCAACCGCACCAATTTAAAGCGGTGTAT 2700
2820 gtctcgaggaatgacagagcgagcgaggtgtgtccagctgtactccacgacggtac 2879
2701 ATTATCGAAGATGCTGTCTGTGGAGATGAGGTGTCAAAATTTATGTAAGATTTGGAT 2760
2880 agagacgaggtgtacacacaggtacgatacgaatgaaacttgcggtgaaccccgactg 2939
1684 ACATTAAATGGGAAAGCGAATATATCTCTCGATAGACAAAGACACAGGAGACATTCAT 1743
1860 gtgagcactggaagatgacctctgactacacacgacagaaacactcttggttgaagtgcc 1919
1744 gtcacgacgagacgactcttcttgatcattcaccgaaaggaactcttggttgacgacga 1803
1920 gctgcagatcttggcgagcgccacacacacagctgacccaccgctggtagaa 1979
1804 gctgacgacacccctgggga---ACCATTCACACGGCGACTTCACAGTCTGATCCAT 1860
1980 cttgagagatgtaacaacaactctccacccctacgcttgccccgttcgactccaaagctc 2039
1861 ctgacgagatATAAACAACCCCTCTACTGTGAGATTGCCGGGAGGCTCTATATGTC 1920
2040 gagggagacgctcccggaagatacagagatcccgggaatacctgctacgacccggac 2099
1921 GAAGGAATGTCCTCCGAGGCTACATCAACAGTAGAGTCCGAGCCACCGATCCTGAC 1980
2100 accagcgactctgtgtgagatcgactggaactcaccactcggccacacagagagc 2159
1981 ACTAGGCTGAAGCTGCGCTTCGAAATCGACTGGACCACTTCTACGCAACCAAGAGGC 2040
2160 agagagacacacactactgaaatcagtcggtgtagatgatacgaacgataccacc 2219
2041 AGGGAAGCTAATCTATTGAATTCACAATTGTGCGAGATGGAACAATATATCTCTGG 2100
2220 gagggacacccgggtctccgacatcgggcgctcgtgtgtaagagataccgggaacgctc 2279
2101 ATCAACAACCGGTGGCTCGCATATAGGACGCCCTCGTCGTAACAAAATAAGGAGAGATC 2160
2280 accatcgactcgaggaattcgaatcgatgcttaccctcaccgctcgctgaggaactcacc 2339
2161 ACAATTGACTAGAGAGAGTTCAGAGTGTGTATCTGACGGTACAGTGGGACCTTTAAC 2220
2340 actgctacgagatgactgaaatcagatgagggcgacgcttcacgatacaataatcagatgac 2399
2221 ACTGTTATCGGGGATGATTAATGATGATCAACCTTCACCATCAACATCATGATATGAAAC 2280
2400 gacaaacgacgactcttcggaacgagcagctgacgacgagatgacgctgacgagctg 2459
2281 GATTAATCTCCGATGAGGTCCCGGGAACACTGGAGCAGTGGCTCGTTCGGGAAATG 2340
2460 gcgagcagcgagacgctatcgctcgtctcgacacgacatcggagcccgctcacc 2519
2341 TCAGAGCGCGGCTGCTGATTTGGGACTCTGACGGCACTGACATCGAGCGACCACTTAC 2400
2520 aacaaagtgcgtacacatacaactagaacaacactcccgaggaattagtgagaatc 2579
2401 AATCAAGTCCGATACACTATGAAAGCAACGAAGCACACAGAGAAATCTATGTGATC 2460
2580 gactcacaactgtaaatgtaggtgtagacgaaagcgatcgtatgcagacgaaccc 2639
2461 GACTTCTACACTGGTCAATCAACAGTGAACAAAGCGCGCAATAGATGATGATGTTCCA 2520
2640 tggcgctctactgttactaacacgctcactgctgacgagctgctccctggaacccgc 2699
2521 CGGAATATCAACCTTACTACTGACTGTGGTGCACCTGATCGTACGCTGAACACCA 2580
2700 aacgaatgtctccacgactcgaactcgaactcgaagtccagcgagatataagaaatgac 2759
2581 GACGATTCCTCGAGACGACCAACTTATGGAAGATCCCGGCAAGTGTGATACAAATC 2640
2760 atcgacaacaacaagaatgacctgagcgctcactgagaaatgtaacagacggtgtac 2819
2641 ATAGTATACCAACAAATAATACCTCAACCTCAACCGCACCAATTTAAAGCGGTGTAT 2700
2820 gtctcgaggaatgacagagcgagcgaggtgtgtccagctgtactccacgacggtac 2879
2701 ATTATCGAAGATGCTGTCTGTGGAGATGAGGTGTCAAAATTTATGTAAGATTTGGAT 2760
2880 agagacgaggtgtacacacaggtacgatacgaatgaaacttgcggtgaaccccgactg 2939
2761 AGAGATGATATTTACCAACAATTCGTTACCAAAATCAACTACCGCGGTGATCTGCTG 2820
2940 cgggattcttcgaggtgagctgacacactgctgctgaggtgacatcaccggggac 2999
2821 CGGACTTCTTCGCTGTTGATCCGGATACGGGAAGATGATGCTACTACCAACCCAGC 2880
3000 gaaaatctgagccgagatgggtagatgacctacacatactatcttgaatcactgac 3059
2881 GAGGTCTTCAGCCGGAGGAGGAGACCACTCAACACAGATCTTCTTAACCTCATTCAT 2940
3060 aactcttctctgactgtagcgttagaggaacacaggaacgaagtgtgaatatgtgct 3119
2941 AACTTTTCCAGCAAGAGACGGCAACAAACCAAAATGATGCCGACTGTGCTGTA 3000
3120 ctattgtagtgaagacaaagctctctgagatgccattgctgtatgaaactcgggttat 3179
3001 CTTCCTTGAAGGAAACGAACGATCTCCAGAGCTCCAGAACCCGATGAACTGCGGTGCG 3060
3180 gttccgagagagacgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 3239
3061 GTGTGAAAGTTTAAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTAAGGTA 3120
3240 gatgacacagacagacgaactcgcgtgctgctgctgctgctgctgctgctgctgctgctgct 3299
3121 GACGAACCTGACAGGAGACCACTCTGGGTGCGGTACGATCATCATGAGCTCAGCATGCGC 3180
3300 gacgagacatcgaggtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 3359
3181 AACAGGAATTCGAAGTCCAGAGCTTTTCAACATGATGATGATGATGATGATGATGATGAT 3240
3360 gaacttgaagcagctatgacttgaatgaggtgaggtgaggtgaggtgaggtgaggtgaggt 3419
3241 GAACGTGAACCCGATGACTGAGAGGCTATTGAGGAACTTATGCTATCATATTTAA 3300
3420 gctctgacacacgctacacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 3479
3301 GCCTACGACCAACGGGATCCGCAACAGATGTCATGAAAGTGAAGAGTGAAGTGAAGTGA 3360
3480 cctacacactccacacacactgctgctgctgctgctgctgctgctgctgctgctgctgctgct 3539
3361 CTTTACAACCTTCAACGCTCTGTTTCTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3420
3540 tctagggagcgcggaacgagcgctgctgctgctgctgctgctgctgctgctgctgctgctgct 3599
3421 GCCAGGAGAGGCTGCTGTCACAGGTTGCTGTCACAGGTAAGAGGTAAGGTAAGGTAAG 3480
3600 ccgactcagccacgacgagacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 3659
3481 CGCATCGTCCGTACCGATGAGAGAGGCTTACATGCTGAGACAAATGCTTGCATGAGTGGT 3540
3660 ggaatgagagcgctcagctgacttgaataaactggaagtggagacgaggaataagc 3719
3541 GGTGATACCTGAAGCCGTCGATTTATTTCCACATV-----GTCAACGACGAGGAGAACGT 3594
3720 gggcagcttataatgacgacgacttccagagcaaatcagacaatcagatacagatc 3779
3595 GGCACCTGTGATGTGAAGCAACCTTCCCGAGGACATTAAGAAATTTGAGTGATCAGATA 3654
3780 cgggcacgagcgcgagacgagacgacgacgacgacgacgacgacgacgacgacgacgacgacgac 3839
3655 CGTCTACTGACGGTGGGACGAACGAAGCCGCTTCACAGGATGCTGACCTTCAAGTGTG 3714
3840 gtctcgtacacacagggcgacccagtgctcagcgaanaatgacgactgctgctc 3899
3715 GTGTTCCGCGCATTAAGGAAGCACTATCTCCCTACGTCGACCTCACACGTTGCTTTC 3774
3900 ttgagaggtgaaagagcctcctcgtagaggtttgaagtcgacgagagagagacttaaa 3959
3775 ATTGAAAAAGAGCGGCGCTTTTGAAGCTCATGAGCTGCGGCTGTGAGGATCGGAAA 3834
3960 aaccacactctcggaagatgactgccaagatatactactacaggttattatgacggaacac 4019
3835 AACCATCTTTGCACTGATGATGATGTCACAAATATCTACTACAGGAGATGATGAAATAT 3894
```



```

EEVYIRATDGTGTEPRPLSTDCSFVVFPIQGEPIPTSTHVALEKEAGLEERHE
LPRAEDRKNHLCDSDCHNIYYRIVDGNNDQPHVDEIDTNYEVLKEIDRSENYTLF
IAANSNPGTALTSTITITVYNREADPOPEVFDISTDSTINRELLITIOATH
SENNPIYITIDLSMTVDPTLASVRETFILNPHTGVTLTIOTPAASHHMFEOVA
TDPAGYSDRANVKIYLLISTNRNVEFLEVNTEQVEQNDLTAQFSAEFEMTCNDVOY
VPADASGVIMNGITEVNGHFIKRDNPVPABEITLRODMWLIAIOISTATRLVLR
DLFTDTSAPADGASAAVYALAVALSALLACLILVFIIRFKLNRLRALTVKY
GSVSGLRNRVGIABGTNRKHAVEGNSNPINMETIKAPDPDSMSDASNDLIGIDLPK
FGENNYPRDDEKTDKPEDIVATHNNNFKSTPSPERANOPKL"

misc-feature
/feature="bcr175b"
/feature="putative
transmembrane region"
5155..5172
primer_bind
/feature="bcr175b"
/feature="PCR primer site"
BASE COUNT 1482 a 1295 c 1224 g 1171 t
ORIGIN

Query Match 35.6%; Score 1959.2; DB 3; Length 5172;
Best Local Similarity 63.0%; Pred. No. 0;
Matches 3197; Conservative 0; Mismatches 1833; Indels 48; Gaps 9;

Oy 144 aactaagatcatgcaaaatgggggtgaaggtctctcccgacagtgctgctct 203
Db 7 AACAAAGGACATGCGAAGTGGAGTTGACGTTCCGATTTGACATCTTCTTAATC 66
Oy 204 ttgacctgcccgaactgcaacacgaagtgctgcatatcatgcaatcaaacagccg 263
Db 67 TATGCGAAGAACTGATTTGGCTCAGAAAGATCGGGCTTCATGTTGCCATACCGACCA 126
Oy 264 gagactccggaactgcccctcttgatcaagaagaaatcatgagtggaacgacctta 323
Db 127 CCAGGCGGAGATCTCCGTAAGCTTGAAGTCAAACTGTAAGGTCGTAAGGACACATA 186
Oy 324 ataccgcccgcgaacgaggaagtatgtagagaact-----tctaccgagat 374
Db 187 ATACCGGGCGGACGACCGTGAAGATGTCTGATGACGGCTACACCGCATGACACCTTACC 246
Oy 375 caaatgaticaggtcatatatacatgaggaagaagaatcgaaggaagctcatcatgag 434
Db 247 TACGGAACCTGAGATTATTTTACATGGAAGAGATGAGAAGAGAGTACCCATCGCTAAA 306
Oy 435 cttaactatcaaggttcaaacagcc-----gtgctgctgcatatgccaagccagcc 488
Db 307 CTCATTTACCGGCTGCCAAATGTACATATGAGCCAGCATTTCTGTGTGTTCTTTTC 366
Oy 489 agagccagcgtgcccgtgaggttgcagacgaatgaagcagagcccatgagacctgtt 548
Db 367 AATCTACTGTCGCCGTTATCCGACAGATACCAGATTAGTAATGGAGAAATGGACCTTAT 426
Oy 549 attacgcaaaagcaagactacgacgaacacacatgacagctatgcttctcaatccaa 608
Db 427 ATTACTGACAGACAGATTACGAGACTCCGGGAATGACAGATAGTGTTTTCAACATCAG 486
Oy 609 gtggaggtgtaatacaagccgtactggtggtggtgagatgtaatacatgcaagaacat 668
Db 487 ATAGATGGGGGAGACATTTGTCGGGTTTTCATGCTGATTTGTAATAATGACCAAC 546
Oy 669 ccggccatccctcaagtgctcagcgctgcgttaattcagaacacagcgagagctagct 728
Db 547 GCACCTTATATTCAGACACTTCAACCGTGTCAAGTTGACGAACCTAGGGGAGAGCTGTG 606
Oy 729 accgactgcgtgtacaaagtgtcagaacgcgatgaatgaatcaagcagccgtcttca 788
Db 607 ACTGATGCGTATATGCTGTAACTGATCCGATGCGCCATCAGCACCAGTCTATGAG 666
Oy 789 ttccgtgtcgaagcagcagcggtgtcgaatgaagaacatcttcaatggttggagaatac 848
Db 667 TTTAGATTCGACGAGATCG--AGGAGACGACAGATCTTCTTACATCCAAAGGGCGAAG 723
Oy 849 gaccacagcagctgttcaatatagaagtgaactgtggtgagatcaatccgacctga 908

```

```

Db 724 ATTCGGGGGAGATGATCGGATGACATGACCGTTGGCATTAATGAACCGTTGAATTTC 783
Oy 909 gagacaactcaagctcatatattagcgtlcaacagcttctgactgcgtcaccgaacaac 968
Db 784 GAGACTTAATCCACTACATATATTTCTGTGACCGCTTGGAGACTCTCTCCAAACCTGAC 843
Oy 969 acggtcacacatgattggtgaagtggaagcgtgagtgcttgccccctccgctgggtgag 1028
Db 844 ACTGTAAAGCTTATGTGTAACATGATGAGATGTGAAACATCGGCCCGCCACGATGGGTAA 903
Oy 1029 atcttcaagtcagcagcttgacgaagaagataatcaagaagcttccctccgcgcagata 1088
Db 904 ATCTTGTGCTGTACACATTCGATGATGAGAAACTGCTCAGAGCTTCCCTGTAGAGCTATC 963
Oy 1089 gacggggaacagcgggaatcaatagggccatcaactataccctcaatcaggatgacgtgac 1148
Db 964 GACGGAGATACAGGCATCAACAAACCTATCCACTACAGCTGGAAGAGAGAGAGAT 1023
Oy 1149 gacttctt-----tccctgaggtgattgaagacggaagctatctcgaagtgaat 1199
Db 1024 ACGTTTTCACATTCGACACCATCGAGGGGGCTGTACCGGCGCATATTTATACGTAGAT 1083
Oy 1200 gagatcgacccggaacagcttgaagaagagcttccaacctcaacctgcgttgcttaacaa 1259
Db 1084 CCAATAGACAGGAGACATTTACAAAGAGAGAGTTCATTTATCCATTAATCGCTATATA 1143
Oy 1260 tctactgacgttagcttgcacaagagggcccaatttcatcatcgtlcaacgagctcaat 1319
Db 1144 TACGACAGAGATCTTACGCTACGCGGCAACGTTGCATTAATGATGATGATTAAC 1203
Oy 1320 gatcgagaccccgagccgctgcatataagaatagatgatatcatgtagaggaactcca 1379
Db 1204 GACCAAGACGAGAACCTTTGTCAAGAGATATCGTTGAACATTAATGATGAGAGACACT 1263
Oy 1380 atgacttcaactcaatgaagaatlttgatccaatgacagatttggtagaagaagct 1439
Db 1284 TTGACGTTGAACTTGATGATCAGGAATTCGATTCATGATAGAGATCTGGGTCAAAACCA 1323
Oy 1440 caatacagagtggaacttgagagagctgttccgcgaagggcggtgccgcatcttcatatc 1499
Db 1324 CAATACACCGTAGAGCTTAGAGAGCGATTATCCAGACAGCGGGCGCAAAAGCCCTTTATTA 1383
Oy 1500 ggcgcgggagcgctcaccagaagagacacctcatcatatgggacacataacacacatg 1559
Db 1384 GCACCCGAAAGTTGGCTACACAGCAGACCTTCATCAATGAGGAGACGCTTAACCTAAATG 1443
Oy 1560 ctgagattcgaagatgctatcttcaagaacatcatcatatgaaggtcaagcgagtgagata 1619
Db 1444 CTCGATTTATGAAGTTCCAGAAATTTTAAACGGATCAGGTTGCCGGTAAATGCAACAGACAT 1503
Oy 1620 aacacgctacgacagctggtgcgagcgctggtgtaacgtgaacgtlcaactggaagcag 1679
Db 1504 GACATTAAGAAACAGCTGGGGTTGCTGATGTATCATCAATCTTATCACTGGAATGAT 1563
Oy 1680 gaacttccatcttcgaagagagcagctaccgccgtgctgttgaaggaacgctgcgcgc 1739
Db 1564 GAAAGCCCTATTTTGAACATATGCGCTACAAAACGTTTCTTCAAGAAATGAAAGGAG 1623
Oy 1740 ggcctcccggtggtgcgaaggtgctgcctccctcgacagaagacatcgaagagtgtagtat 1799
Db 1624 GGATTTTTCGAGGCTAATATGTAAGAGCCACAGACAGACATTTATGAGAGATGTAACAC 1683
Oy 1800 tcatgatggcaacagcgtgttgaactcgtlcatagatagatgaatcaagggagaagatcttc 1859
Db 1684 ACATTAATAGGGAAGACGCAAAATTAATCTCTGATATGACAGACACAGGGGAGATTCAT 1743
Oy 1860 gtgagcagtcagcagcttgcctgacacacacggaagaaacactcatattggtcaggtgc 1919
Db 1744 GTCACGACAGAGAGACTTCTTGTGACTATACCGACAAACGAACTTTTGTTCAGGTCAGA 1803
Oy 1920 gctacgatacttgggagagcgccacacacacagtgacacacccagctgtgtagaa 1979

```


OY	4140	gggagccctctgagagacttccatccctcgttgttaagtgtaatgtaagaaagcaaccg	4139
Db	4009	GGCGGAATCGCTCTTAACCTCTTACAAATAACCATACCGTTAAATGTGAAGAGCAATCC	4068
OY	4200	cgcccaatttctcgagcaagaccttcaacaagcgggacttgcagcttgagcaagcttgc	4259
Db	4069	CAGCCGTACTTCGTCACAAAGATTGTGTACACGGCTGGGATCTCCACTTCGGATTCAATCAAC	4128
OY	4260	cgggaattgtcttactgtctcaaggcgagcgcaacaagaagacagacatcaagtaacca	4319
Db	4129	AGAGAACTGTTGATTATTACAGCGGACACATTTGAGAAACGCAACCGATATATCAACGATA	4188
OY	4320	gaccgtgcgagcatgcagcgtcgagcaagcagcctagaagccgctgcgcgactcggcttcgcg	4379
Db	4189	GACTGTGTCAGAGATGTAACGTATCCTTACCTTCCTGCTCCTGATAGGAGAGACACCTTACTATA	4248
OY	4380	ctgcattgcgaccacacggcgctgcttccgtccaatataatgacccacacgcttccatgcagcgc	4439
Db	4249	CTTAACCCACATACCGGTTGTAAGTACTCTTAACATCCAGCCCACTGCTCAATGACGGA	4308
OY	4440	atgttcgagttcgacgctacatcgcttaagatacagcttctgtcaatcgcagacagcccgctgtg	4499
Db	4309	ATGTTTCGAATTCAAAGTCGAGGACAGGATCCAGTGTGGCTACAGTATGATGACAAATGTC	4368
OY	4500	aaagttactactccatccatcgcgcaaaacgcgctgcacttcaatttgcataaccaattgcg	4559
Db	4369	AAGATCTACTTGATCTCCACAAAGGAACCGGGTCTTCTTCTTTGTGTTAACCTTTAATAA	4428
OY	4560	accgtttagcaagaacagaaatttccatagaagcgcaagcttgaagacgggttcaacatagc	4619
Db	4429	CAAGTCGACGACGAAACCTGACTTCATTAAGCAAAACGTTTCCGGCGGCTTGCAATGACG	4488
OY	4620	tgcaacatcgcagcagcagtgctgcgcttccagcgagcagcagcgcgctgtgcgcaagacagacac	4679
Db	4489	TGTAAATCATCGAACAAAGTCGAGCGGGCACATGACGCGAGGGGGTCAATGAACGGCATC	4548
OY	4680	acgagagtgctgcgcgcaacttccatccggagacaagctgtcccggtgtcagagcaaaagtgctag	4739
Db	4549	ACCGAAGTGAAGGGGACACTTCATCCCGGATTAATGTGCCGTGACTCCGACGACGATTCGAG	4608
OY	4740	gcgcgtccgcaagcagcagcgtgcgtctgcgcgacatccatccatgactgactgaagcaacaagc	4799
Db	4609	ACTCTCGAGGGGACATGAGTGTGCTGACTGCCATCAATCCACGCTGGCGACGGCGCTG	4668
OY	4800	cttggctgtcgaagacccctgtgcagcgggtgacactcgcagcgtcagcgagtgcaatgctag	4859
Db	4669	CTAATGCTGGCGGACGCTTTCACGAGAACACAGTCGGCGGCCGACGCGGGCTC---GGCC	4725
OY	4860	atgcgcgtctacgacatacgcgcgctctccgcgtgtgtactagcttccctctgcctgtactg	4919
Db	4726	GGGGTGGTGAAGCGGCTGGGCGGTCTCTCAGCCCTTCGGCTGGCTGTGTCCTACTAG	4785
OY	4920	cttcctgcattgtctctgtagaagaagaacactgcgaacggcagctgcagaagctcccatg	4979
Db	4786	CTAATCATCTTTCATCATCAAGAGACTTAATAAACTTAACCGGTCTTTGGAAGCCCTGAACGGTT	4845
OY	4980	acgaagctacggtctcgtgcgactccgggctgcgaacccgcgcggcgt---ggcgccggagcaac	5036
Db	4846	AAGAAGTACGGATGGTGGATTCAAGTTTGAACCGCGGTGGGATCGCGCGCGCGGACAC	4905
OY	5037	aacaagcagcgcgtcgcagggctccaaaccccatgttgaaagagagccatccgcgcgcgac	5096
Db	4906	AATTAACATGTCCGTGGGAAGGCTCCAAATCCGATCTGGAAAGAACACATTAAAGCCCGGAC	4965
OY	5097	ttcgaagccatcagttgacgcgagtgtagcgactccgcagctgcatctgcatagagactgcg	5166
Db	4966	TTTCATTTCTAAGACCGATGATTAACAAAGCATCTGATCTTAATGGCATTTGAAGATCTGCCA	5025
OY	5157	caattccgcgagactacttccgcgcgcgcgagcaaga	5194
Db	5026	CAC TTCGGCGAAATATATTATTTTCTCGGTGACGTGCA	5063

[illegible]

Db	3058	GGTGTGGACCGGTGATGTGATGACCAACGACCGATCTTTTAAACCTATGACAA	3117
Qy	3062	ctctctctctgtagtgtagagtagagaaacagagacgaagtgtgaataatctgtcgtctc	3121
Db	3118	CTTTCATGGGGGAGAGAGAGTATACAAATATCAAGAACAGCAAGTCTCGTTATCT	3177
Qy	3122	attggaagtgaagcaacaacgctctctgagatgcatctgtcgtatgaaatctcggttgtatgt	3181
Db	3178	GTTCGATGTCGATGATGCATATGCTCTCTGAATTTGCCACGCGCAGCACTCTCTTGACATAT	3237
Qy	3182	ttccgaagagcagctgtcgtgtgtgtcgtgtacccccaagaatctaacgacggagcaagga	3247
Db	3238	ATCTGAGAACCTTAAGCAGGGCGTCGCTCTTGAAACCAATATCTTGCCCCGAGCCGCA	3297
Qy	3242	tgaaccaaagcaggaacaactcgtcgtctcgtgttaacgaatcttgacactcaagatcaagga	3307
Db	3298	CGAGCCGACACACAGACAACCTCCAGGTCGGCTACGAGATCTGTAACCTC--AGCACGA	3355
Qy	3302	ccgagacatcgagtgccggaatctctaccatgatctcogatgtgaaacaaacacacgtggga	3367
Db	3355	GGGGGACATTCAGAGTCGCGAGCTGTTTGATGATACAGATGCGCAACGTCACGGGAGA	3414
Qy	3362	acttgaagccgtctgagactgagggggtatttgggacactcgaacaaatcttcattgagcc	3422
Db	3415	GCTGGAGACCGCCATGAGCTCAAGGAGATTTGGGGACGTACGTATACATATCAGGGCC	3474
Qy	3422	cttcgacacagcgttaccgcgacgagagtgccacgagacgtatacccttgatcacgccc	3481
Db	3475	ATTTCAGCACGGCATTTCCGCAATATGTCATGAAACGACATATAGACTGATATCATCATTCC	3533
Qy	3482	ctacaacttccacaacccctgtgtctgtctgtctccgcgaaccccgactccgtatccgcttc	3541
Db	3535	GTTCACATCTACACCGCTGCTGTCTCTCCGACCAACGATGCCCCGTATACGACTTGC	3594
Qy	3542	taggaagcgacgaacaggaagcgccctctcggagaaagcgtgcgaacagttctctgaacc	3604
Db	3595	GAGGACAGACCTGTATTCATATGAGATTCTAGCGACATGTAACGAGATCTTGAGACG	3654
Qy	3602	gattcaagccacgacagagagagcgtcccaacgcgagacgctcaacgtctcaagc	3661
Db	3655	GATATCGCGCATGATCCGGACGGACATCCAGCGGGCGTCTACCTTCCAGTGCTAGAG	3714
Qy	3662	aaatgaggaagcgctctcaagtattgatatataactlgaagtggagcagagaaataagcgg	3721
Db	3715	CGATAGGAAATCAACAAGGTAATCTTCAAGTAGTAACGATGCG-----GAGAACCTCGG	3768
Qy	3722	gcagcttatataagcgaagcttttcccaagggcaatctcagacaattcagatacgaatccg	3781
Db	3769	CTCGTTGAGTTACTGTGAAAGCCGTTCCAGAGGATCAAGGAGTTCCGGATTAACGATTCCG	3828
Qy	3782	ggccaaggaagcgagacgaagcccgccgctttgagccgaagctcaacgttttgggtgt	3841
Db	3829	CGCTACAGAACCAAGGAAACGACCCAGGACCGCTGTCCACGGACATGACCTTACAGTTGT	3888
Qy	3842	cttcgtaccacaacagggcggaaccagtgltcagcggaaaaatgtagctactgtctgcctctc	3901
Db	3889	TTTTGTGCCCCGCAAGAGAAACCTATGATTCGGCTCCTCAGAAATGCTGTGGCTTCAT	3948
Qy	3902	cgaaggggtgaagaagcgctcgtgtgagagtttgaagctgtccgcaagcaagagaccttaaaa	3961
Db	3949	AGAAAAGATGCCCGCATGGAAGATGTCACCAACTTCTCTAGACAAAGATCAACAA	4008
Qy	3962	ccaactctgcgaaatgaatgcagcaagatatctacatacagtgattctgacggaanaaaga	4021
Db	4009	CCATCTCTGTGAAGACACATCTGTACACGCAATTTACTATCTATTTACGATGGCAACGGA	4068
Qy	4022	gggtcttctgtaactggaacagtcgaacgaacgtatcatctcccttgcgcagagattgagccg	4081
Db	4069	AGGTCAATTTGGGCTCGATCTCTTGGCAACAGGTTGTTCTCTGAAGAAAGACGTATTAAG	4128
Qy	4082	cgaagtgagcagcgtcttaacgctgcacatcgcgcgagcaactcgcgcgaagcgaatcgg	4141

Db	4129	GGAAACAAGTGGCTCCCACTACTGCAAGTGGGGCTAGTAACTGGCCGATG---	GTGG	4189
Qy	4142	gattccctctgcagacttccalcctcgttgtcaagctcaatgtlaagaagaacgcaaccgcg		4201
Db	4186	CAATTCACATTCCTGGTTCCATTCCTTACTGTCACGTATTACCGTAGAGGAGGACACCTCG		4245
Qy	4202	cccaattcttcgaagcaacttacaagcgggcaatttcgaagtttgacagcatgtgcg		4261
Db	4246	TCCACTGTTTGTGAGGGAAATTGTACACCGGAGGAAATCCACAGCGGACTTCATGGGAG		4305
Qy	4262	ggaatttgtttctctgaagggcgagccacacagaagaacgacacataacgtatacacatata		4321
Db	4306	AGAGCTGCTCAGANTTACATGCGAGCCACGTCAGTGTGAAGCTGGCCATTACTTATGATATGA		4365
Qy	4322	ccgtgcgagcaltgcagcttgcagacagaccctagaagccgtgcgcgactcgcgtctgcgc		4381
Db	4366	CTACGATACATGGTAGTGGAGCCCACTCGTAGGAGCATGACAGTGGCGCTTTCGTACT		4425
Qy	4382	gcattgcagaccgcggcgtgtcttcgtctcaatatgcagcccacgcgttccatgcagcgtat		4441
Db	4426	GAACCTCAAAACCGAGTGCAGAGCTTATATCCAGCCACCGCCAGCATGATCAAGACT		4485
Qy	4442	gttcgaattgcagctcaatgcgtatgaataagttctgcaatcgcagacaagcccgtgtgaa		4501
Db	4486	GTTCAAATTTCAGATCACAGCTATGACAGCGCGCGCTCAAGACCCGACCGAGGTAC		4545
Qy	4502	agttaccctcatctcatcgcgaanaaccgcgtgcacttcatttcgataaaccacattgcagac		4561
Db	4546	CGTGAAGTGGTATTCCTGGGAACCGCGCTACTCTCGTGGTTCGTCAACAGCTGCACMA		4605
Qy	4562	cgttgcagacagaagaattcatatagcgcgcacggttcagacaccgggttcaacatgcgt		4621
Db	4606	GGTCGACAAACAAGAGACTTTATCGGGGACACTTCAACGGTGGGTTTCATCATACTCG		4665
Qy	4622	caacatgcagcaagttgtgtccggttcaagcga---cagcagcggcgttgcgcgaagaagcac		4678
Db	4666	CAACATTCGACCAAGTGTGTGCCCTTACAGCACCCGCTACACCGGCTGGCGCTGGACACAG		4725
Qy	4679	caccgaagttgcgcgcgcgaactlcatcgcgggaacaagctgcgcgtgcagagcaagaagttcga		4738
Db	4726	CACGCAAGATG-GGGGCCACTTTCATACGGGACAAAGTACCCTGATACGCGATGATATGA		4784
Qy	4739	ggccgtccgcagcgacaacggtgtcgtcgcacacatccagctgcgtatgcvtgcacaacag		4798
Db	4785	-CAGATCCGAGTACCTTACCTCTGACCTCGATACAAACAAACGCTG-GGCGGCGAT		4842
Qy	4799	ccgttgtctgaagaacctgttgcagcggtgcacatccgaagcgtcaggcggaggtcaatgca		4858
Db	4843	CGTGCTGTTTGGAGGACTTGTGTGACCACTCCACCCCGGACTTGGCGCTCATCTGAC-		4901
Qy	4859	gattccgcttcaagcactaagcgcgcgtctcgcgtgtgcataagcttccctcgtcgtact		4918
Db	4902	-CTGACCGGTACGTCGTGGCTCCTACTGCTGTGTGCTGGTTTCAATGTGCTGTGTCT		4960
Qy	4919	gcttctgcgcatlgtctctgtagaacaagagcactgcagcgcgcagctgcagcacttccat		4978
Db	4961	ACTGCTTACTTCATCATCAGACGTAGAGGGCTTAACCGACAGCTGTGAAGCCCTGTGCAT		5020
Qy	4979	gacgaagtaagcgtccgltgcatacccggtctgaaacgcgcgcggct---ggccggcgac		5035
Db	5021	GACGAAGTACGGCTCACTGTGACTCTGGATTGAACCGCGCGGCATCGCCGCCGGGAC		5080
Qy	5036	caacaagaacgcgcgtgcgaaggtccaaaccacatgtgaaacggaggtccatccgcgcgcga		5095
Db	5081	CAACAAACACACTGTGGAGGCTTCCAACTTCTTCAATGAAGCAATTAAGACCCAGA		5140
Qy	5096	cttcgaagcatcaagltgcagcgatgggcagactcgcagctgcgtcatcgcagatgcac		5155
Db	5141	TTTGAATGCCATTTAGGAGGGCTTCAACAGACTCTATCTGATCGGCATCGAAGATCTTGG		5200
Qy	5156	gcaattccgcgaagcactacttccgcgcgcgcagacaga		5194
Db	5201	GCACCTTGGCAAGCTTTCATGTGATCTCTAGGTGAACA		5239

RESULT 8
LOCUS 177078 177078 5577 bp DNA PAT 03-APR-1998
DEFINITION Sequence 1 from patent US 5693491.
ACCESSION 177078
VERSION 177078.1 GI:3013232
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5577)
AUTHORS Bulla, L.A. and J.L.T.
TITLE Receptor for a Bacillus thuringiensis toxin
JOURNAL Patent: US 5693491-A 1 02-DEC-1997;
FEATURES
source 1..5577
BASE COUNT 1476 a 1496 c 1404 g 1201 t
Query Match 35.6%; Score 1954.6; DB 6; Length 5577;
Best Local Similarity 63.9%; Pred. No. 0;
Matches 3248; Conservative 0; Mismatches 1769; Indels 62; Gaps 17;

844 CAGCGTTCG---TGCGGACGAGGAGACCTTCTACATCGAAGCGAGCAATATCCCAACCA 900
QY 860 ctggttcaatatagaagatgactggtgggatacttcgcccctgaacttgagacaactca 919
Db 901 ATGAGATGGCTTAATATATGACATAGCGCTTATATACCTGCTCAACTTGTCACCGCTCC 960
QY 920 gcttcaataatlaagagctgacagctctctgactcgcgtaccgaacaacacagcgtaccat 979
Db 961 GCTGCATATATTCAGGCTACACAGCCCTGAGCTCGCTCCGGAACACCCACAGCGTACTAT 1020
QY 980 gatgtgtcaagtgtggaagacgttagagcttcggccccctgcgtggtgtggagaacttcctcagt 1039
Db 1021 GATGTGTCAAGTGTGCGAAGTGAACAGCCGCTCGCCCTGCTGAGATCTTCGCTCT 1080
QY 1040 gaagcaggttgaagagaaagaaactaatcagagcttcccccgcgcgcgatatgagggagac 1099
Db 1081 CCAACAGTTTGAAGAGAAATCTTACCAAAACTTCACAGTGAAGGCGCATGACGAGAACAC 1140
QY 1100 ggaatacaataaggccataactataccctcacaaggaatgaacgtcgaagactcttcc 1159
Db 1141 TGAGATCAATATGCCATATACACTACAGGCTGATCAACAAATGAGGAAGACATTTCTTACG 1200
QY 1160 cctbgaagtgatt-----gaagacggagctatctcgaagtgaactgagatcgacg 1210
Db 1201 CATTTGAGGCCCTGCGCTGATGGAAGGCGGCGCTGATTCCTCGTGCCTCAATTGACCG 1260
QY 1211 cgacaagcttgaagagagcttcttaacctcaacatcggttgcttgaatactcagagc 1270
Db 1261 CGAACACTGTCAACGAGAGGTGTTCCACTTACGATCGCTTCAAAATATGATGAGGA 1320
QY 1271 tagcttgcacaagaagccacatlttcaactcgtcacaacgcgtcaatgatcaagcagc 1330
Db 1321 GGCGTCTCCACATCAACAAAGGTGATCATTTGACAGACATCAACAGCAAAAGACC 1380
QY 1331 cgaagccgctgcataaagaatacaagatgataatcagaggaagaaactcacaatgaactctaa 1390
Db 1381 TGAACCTATACAAAGAAATATCGACTGGCAATCATGAGGAGAGCGCCCTGACCTCA 1440
QY 1391 ctccaatgaagaatttggatctcaatgatcgaatttgggtggaagaagcctcaatacagct 1450
Db 1441 CTTTCGTAAGAATTCGATTTTCATGATTAAGATTTAGCTCAAAACGCTCAGTACAGCGT 1500
QY 1451 ggaacttgaagacgtgttcccgccaagggcggtccgacttcaatactgcgcggggag 1510
Db 1501 GCGTTAGAGAGCGGTGACCCCTCAAGCGCTGAGGATTTCTACATAGCCCTCAAGT 1560
QY 1511 cggctacccaagagcgagacttcaatacgtgggacacataaacacacacatgctgattaga 1570
Db 1561 CGGCTACAGCGACAGACCTTATCATATGAGGCACTTCATCACTCCATCTCGATTTACGA 1620
QY 1571 agatgcatltttcagaacatcatcaatlaagttcaagcaggtgagacatgaagaacgctag 1630
Db 1621 AGTGCCAGAGTTTTCAGAGTATTCGATTCGGGTGATGAGGACCGACACACAGCACGAG 1680
QY 1631 ccacgtggcgagagcgctgtgtgtagctgaacctgaaacttgaagacgaacattccat 1690
Db 1681 GCACGCGGGCGCGGTGGTTGACATGACCTCATCAATTTGGAACAGATGAGAGCGCAT 1740
QY 1691 cttcagaagagagcagctactccgcgtcggttlaagagagcgcgtcgcgcgtcccggt 1750
Db 1741 CTTTCGAAACAGCGCGGTGCAACACCGTCACTTCGACGAGCTGGAAGGAGGAGGTTCTTCT 1800
QY 1751 ggcacaggtgcgcgccttcagacagagacatgaacgagctgaagtagtgcattcattgag 1810
Db 1801 CGCCAGGCGGTTGCGACACGACAGAGACATCGGGATGTCGTCGAGCATCTTTATTGGG 1860
QY 1811 caagctgttgaactactcgttcatagatgaatcaacagggagagatcttcgtgaagcatgga 1870
Db 1861 TTAAGCTGTTAATCTTCACCATCGACAAACTCAACCGCGCATCCGGCTCTCAGCTTA 1920
QY 1871 cgatgctctcgactacacacagagaaacactatattgttctgaagtggtgcgtcgaagatcc 1930
Db 1921 CGACTCCTTCAACTACATCGAGAAAGTGAATATTGTCGAGGTGCGAGCTACAGACAC 1980

QY	1931	tttgggagcagcgcccaacaacacatgtgacacacccagctcggtgtgatagaactctgagatgct	1930
Db	1981	gcttggggcga---accctttccacagcggcgagctgacagctgggtctatagcattaaaggacat	2037
QY	1991	caacaacactctctccacccctgaagctctggccgcttgacatccaagagctgcgagagagact	2056
Db	2038	caacacacacggccaccacactttacggcgctgctccagggcagctcccccaagtggaggaacact	2097
QY	2051	tcccgaagatacagagatataccgggaaatacactgctaccgaccccggaacaacgaccta	2110
Db	2098	gcctgattggccacgctcatcaccacagaggtttaccgcgcaccgcacccagccacacccagcggcga	2157
QY	2111	ccctggtgtgtaagatgcagctgggaactccacactcgggccaccaagccgagggagagacaa	2170
Db	2158	ttctgggcttcagagataaactgggacaccttcttgcgcacccaagcaaggccgcagcgttaa	2217
QY	2171	ccctactgaatacagctcggggtgtaagttatgcagaaacgatatataccccaacggagcgaaagc	2230
Db	2218	ccccgacagatgtttaggaattggcgggaaattcgagaccatctttcccgagatttaaacacgc	2277
QY	2231	gggttcgcgcacatcgggcgctcgltgtgtgaagaagatccgggaacaactcaacatcagatt	2290
Db	2278	gggactgtgctatccggccgcgctgttagcgccgcaaaacacacacagttacacattagacta	2337
QY	2291	cgaagaaatcgagaatgttttaactcaaccgctcgcggtgagggacctcaaacgtcatatcg	2350
Db	2338	cgagagacttttagaggtcctcttccctcacagagaggggtgctgaccttaaacccgtttacgg	2397
QY	2351	agatgaactacgaatgaaggcgagctgtcaacatcaataatcgaatcgaatgaacgaacaacgccc	2410
Db	2398	agagacactacgacgaatcgatgctcaacaatatactaatcgattatgaagacgaacacggccc	2457
QY	2411	gacttcgcgaacggcaacgcctgacgaacgacatgacgtcggtgcgagctctgacgcccagcg	2470
Db	2458	gggtgggctggagggagacactctggagacgaacttccagatcccgagattgctggcggcgcg	2517
QY	2471	caagctatcgctgcggtgtgtgtccaaacgaacatcgacggcgccgctctcaaaccaagtgcg	2530
Db	2518	gctctggtggggctccctgtggcccgacgacgacatcgacggagaccgctcttccaaacaaagtgcg	2577
QY	2531	ctacactataacaactgaacaacaacactcccgaggattagtgaaatgtgacttacaac	2590
Db	2578	atpacacattttccctcggtgaagaacagataagacactgataatgatgtcattccttccac	2637
QY	2591	tggtcaaatgtgaagtgtgaatgcgaacgaagcgatcgatgcagacgaacccctggcgcttca	2650
Db	2638	gggttcacaaatttccctgtgaacacaaacggcgccatcgacgggtatccttccacacgctttcca	2697
QY	2651	cttgttaactaacccgttcactgtctagcgaagagtgctccctcggaacaacggcaagatggcc	2710
Db	2698	ccttactactatracagtggtgctctgtatgtaacgagtgcttgacagaaagatctccgcatggcc	2757
QY	2711	tcacagatccaactactctggaatgtccagggcgatatacgaaatagaataatcatcagacaacaa	2770
Db	2758	ccctgaccccgacttatatgggaaacccaaggaataatcacaatgccatcaccagacga	2817
QY	2771	caacaagaatgcctggaagccgtctcaactcgaagaatcaaacacgacggtgtacgtctggagaa	2830
Db	2818	caacaaagctcccccagcgggaamaacgactaaagttcgatcacctgctgattattttacgaaa	2877
QY	2831	tgccagagcgcgcgacgaaggtgtgtccagctgtcatcttccacgacggctgacagaagcgagt	2890
Db	2878	cgcaaacccactttagacgaagtggtgtcactctgattagccactgattcttgacaagacgaat	2937
QY	2891	gtaccacacggtatagatatacgaatgaactgtcggttgaaaccccgactcgcggaattctt	2950
Db	2938	atpaccacacgctgactgactacgtcatcaattatgcagagaacccctcgactgatgaactctt	2997
QY	2951	cgaaagtgcagctggaactggtgtgcgcttgaaagtgcattaccgg-----ggagaga	3001
Db	2998	ctccgtgaacgagaaagacggcgcttggtgtaactggaactgtatgagaccacagggtaattggcga	3057

QY	3002	aaatltgacccgcatgttggaatgtgagcttcaacatactatcttgytaatttcatgataa	306
QY	3058	GGTGTGGAGCGGTGATGGTGTATGAACCAACGACCGATTTCTTCAACCTCATGACAA	311
QY	3062	ctctcttctctatgltgtacggtlaagagaacaccgaaggttgaaatttltgtctct	3122
Db	3118	CTTCAATGGGGGAAGAGAGGTAAACAGAAATCAACAAAGACACAGAAAGTTCTCGTTATCTT	3177
QY	3122	atltgagtltgaacgaacacgcctctctgatatgtgcaatttcgcgtatgactccggtltgatl	3181
Db	3178	GTTGGATGTGAAATGACATAGCTCTGTGAATTGGCCACCGCGGAGCACTCTCTTTGGACAT	323
QY	3182	ttccgaaggaagcagltgtcgtgtgtcgtgtactccagaatctacgcacccggaagga	3241
Db	3238	ATCTGAGAACCTTAAAGCAGGGCGGTCCGCTGTGAACCAATATCTTCCGCCGACCGCA	3292
QY	3242	tbaaccagaacccgaacaaactcgtgtctcgtttacacgaatcttgactctgaactgaacga	3301
Db	3298	CGAGCCGACACAGACAACTCCAGGGTGGGTATGAGATCTTAACCTC---AGCACGGA	3355
QY	3302	ccgagacatctgaagtlgtccggaatctcttccacatgatctcgatlgaaacaaactgtggga	3361
Db	3355	GGGGGACATTCGAAGTCCGGAGCGCTTTGTGATGATATACATATGGCAAGTCCAGGGAGAA	3411
QY	3362	acttgaagaccgcatgatacttgaggggatltggggcaacttgcgaatatltcaatgaagc	3421
Db	3415	GCTGAGAGCCCGCATATGACCTCAAGGATATTGGGGAGCTATGCTATACATATACGGGC	3474
QY	3422	cttgcacaacggtctaccgcgaaggtccaaagagatcaacctgtgtatccgcc	3481
Db	3475	ATTTCACACACGCAATTCGGCAAAATGTCATGACAGAACATATGAGCTATGATCATCC	3533
QY	3482	ctaacacttccacaacaccctgtgtgtctgtctccgcgaacccgactccgcatctcgcttc	3541
Db	3535	GTTCACTACTACTACGGCGCTGAGTTTCGTCTTCCGACCAACAGATGCCGTATACGACTTGC	3599
QY	3542	taggaagcgcgcacaacagaagaagcgcgtctctgcgcgaagcgtgcacaacgaatctctgaagcc	3601
Db	3595	GAGGGAAGACGCTGTAAATCAATGAGTATTAGGACAGTAAGTAACGAGAGTCTTGGAGCG	3655
QY	3602	gactacgcacccgaacgaagaagcgtctccacgcgcgcgcgcgtcaegtctccacgtccaaag	3661
Db	3655	GATATCGCGCGACTATCCGAGCGACATCCACGGGGGTGTGTACCTTCCAAGTGGTAGG	3714
QY	3662	aaatgaagagccglttaagtaacttgttataacttgaatgaatgagggaggaataatgaacg	3722
Db	3715	CGATGAGGAATTCACACAGGTATCTTTCAGATGATTAACGATGGC-----GAAACCTCGG	3766
QY	3722	gcagcttatactacgcagacttctccacagaacaaatcagaacaatlcaagatcaacatcgc	3781
Db	3769	CTCGTTGAGGTTACTGCGAAGCGGCTTCCAGAGGAGATCAAGGAGTTCCGGATTAACGATTGC	3822
QY	3782	ggccaaggaacggcggaacgaagaccgcgcgcgcgttltgaccgcagctcaegtlttcggtgct	3841
Db	3829	CGCTACACACACAGGGAACGCGACCCAGGACCGCTGTCCACGGAGATACGATCCTTAAGATGT	3888
QY	3842	cttgcgtcccaacaaagggcggaacccaagtggtctcagcgsaaatgcaagctatgtctgtctctt	3901
Db	3889	TTTTGTGCCCCAGCAAGGAGAACCTTAGATTCGGCTCTCAGACATGCTGTGTGCTTTCAT	3948
QY	3902	cgaggtgtgaagaagcgtccgtgaagatlttgtagctgtgcgcgaacgaacgaagacttaaaaa	3961
Db	3949	AGAAAGAGTCCGGCATATGGAAGAGTGTACACAACTCTCTTAAGACCAAGATATCAAGAA	4008
QY	3962	ccaactctcgaagaatgactgtccaagatatctactcaaggtttatgtacgcgaacaacga	4021
Db	4009	CCATCTCTGTGAAMACACGACTGCACAGCAATTTACTATCTATTATTCAGATGACGACAGCA	4068
QY	4022	gggtcttcttctgtaatgtaacgaatcgaacgaagctcatctcccttgcgcgaaggtgttgaccg	4081
Db	4069	AGGTCAATTTCCGGCTGTGATCTCTTGTGCACCAACGTTGTCTCTGAAGAAAGACTGATTAAG	4128
QY	4082	cgaggtgtgcacgcgtcttaacacgctgtcacatcgcggtgcgaacaaactccgcgaacgaactg	4141

Db 4129 GGAACAAGTCTCCACACTCTGCAAGTGGCGGCTAGTAACTGCGCCGAGT---GTGG 4185
OY 4142 gatccctctgcaagcttccatctcgtgtgcaagtcgaatgaagaagcaaccgcg 4201
Db 4186 CATTCACCTCTTCCGTGCTCATCTACTGCTACTGCTTACCGGAGGAGGCGACCCCTCG 4245
OY 4202 cccattcttcgagcagccttacacagcgagccttccgaagcttggaagcattggccg 4261
Db 4246 TCCAGTGTGTTGAGAGGATTTACACCGGAGATATCCACAGGGAGTCCATGTGGCAG 4305
OY 4262 ggaattgttactgtcagggcgagccacacagaaagcagacacatcaagctacacataga 4321
Db 4306 AGAGCTGCTCAATATACATGAGCCAGCTGAGAGCTGCGCATTTACTATTGATATAGA 4365
OY 4322 ccgtgcgagcagcagcgtcggaagcagccttagaagccgtgtgcgcagctcgccctgcgc 4381
Db 4366 CTACATACAAATGATGAGTGGACCCGAGCCTGGAGGAGTGAAGACAGTGGCTTCTGACT 4425
OY 4382 gcatgagacacagcgctgtcttcgctcactatagcagccacgcgttccatgacgagcat 4441
Db 4426 GAAAGCTCAAAACGGAGTGTCTACCGTTAATATCCAGCCGACGCGCATGCTGACT 4485
OY 4442 gtctgagtcgagcgtcatcgtcagataagcttctgcaatcgacacagcccggtgtgaa 4501
Db 4486 GTTCAAAATTCGAAGTACACAGCTACTGACACGCGCGCTCAGAGCCGACCGAGCTCAC 4545
OY 4502 agtcacccatcatcagcaaacgcgctgacgttccattcattcgaataaccactgtgagac 4561
Db 4546 CCGTACAGTGGTATCTCCAGAACCGCGCTGACTTGGTTCGTCACACGCTGCACACA 4605
OY 4562 cgttagcagcaagaatattcatagcgcagccagcttcaagcagcggttcaacatgagctg 4621
Db 4606 GGTGCAAGACAAACAGAGACTTTATGCGGACACCTTCACGCGTGGTTCAACATGACCTG 4665
OY 4622 caacatcgacacagctgtgtgcgtltaagcga---cagcagcggtgtgtgcgaagacacac 4678
Db 4666 CAACATCGACCAAGTGTGTCGCCGCTAACACCCCGTACACGCGCGTGGCGCTGAGACAG 4725
OY 4679 caacgaggtgtgcgcagccttcatcgcgcaaacgtgtgcgtgcagcgaagaagtgca 4738
Db 4726 CACGCAAGATG-GCGGCCACTTATACGGGACACAGTACCCTGATCTGCTGATGATGATA 4784
OY 4739 ggcctgcgcagcagcagcgtgtgtgtgcgcacacacacacacacacacacacacacac 4798
Db 4785 -CAGATCGGTATGACACTGATCTCTCTGAGCTCGATPACAAACACAGCTG-GCGGGCGCAT 4842
OY 4799 cctgtgtctcgaagacctgtgtgaaggtgaacactcgcagcgtagcggaagtgcaatgca 4858
Db 4843 CCGTGTGTTGCGAGGACTTGTGACCAACTCCAGCCGGACTTCGGCGCTGACTCGAGC- 4901
OY 4859 gatcgccgtctacgac 4918
Db 4902 -CTGCAAGGTTGACTGCTGCGCTCTCACTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4960
OY 4919 gcttcgcgcatgttctgttagaacaagacacacacacacacacacacacacacacacacac 4978
Db 4961 ACTGTTACTTATCATCATCAGACGACTGAAAGCGGCTAAACGAGGTTGGAAGCCCTGTGCAT 5020
OY 4979 gaaggaagtaagcgtcgtgtgac 5035
Db 5021 GACGAGATGAGGCTCTACCTGACTCTGATGTAACCGCGCGGCGCATCGCCGCCGCGCAC 5080
OY 5036 caaagac 5095
Db 5081 CAACAAACACACTGTTGGAGGCTCCAAACCTTATTTCAATGAGAAATTAAGACGCGCACA 5140
OY 5096 ctctgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5155
Db 5141 TTTAATGTCATTTACGAGGTTCCACAGACTCTGATCTGATCGGCGATGGAAGATCTTGC 5200
OY 5156 gcaattcgcgagcacttccgcgccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 5194

Db 5201 GCACCTTGGCAACGCTCTTCATGATGATCTGAGGTGAACGA 5239

RESULT 9
AF367362 5355 bp mRNA INV 12-AUG-2001
LOCUS Heliothis virescens cadherin-like protein mRNA, complete cds.
DEFINITION AF367362
ACCESSION AF367362
VERSION AF367362.1 GI:15149239
KEYWORDS
SOURCE tobacco budworm.
ORGANISM Heliothis virescens
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
Ditrysia; Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE 1 (bases 1 to 5355)
Gahan, L.J., Gould, F. and Heckel, D.G.
AUTHORS Identification of a Gene Associated with Bt Resistance in Heliothis virescens
JOURNAL Science 293 (5531), 857-860 (2001)
PUBMED 11486086
REFERENCE 2 (bases 1 to 5355)
Gahan, L.J., Gould, F. and Heckel, D.G.
AUTHORS Direct Submission
TITLE Submitted (31-MAR-2001) Dept. of Genetics, University of Melbourne, Parkville, Victoria 3052, Australia
FEATURES
source
1..5355
/organism="Heliothis virescens"
/db_xref="taxon:7102"
/note="s1 allele; CryIac-susceptible"
10..5208
/note="CaBP"
/codon_start=1
/product="cadherin-like protein"
/protein_id="AAK8196.1"
/db_xref="GI:15149240"
/translation="MAVDVRLTAAVLILAAHLTVADDCSYMAIPPERPDPNQNF
EGVMSQPLLPADREDVCAADPDSLANLTVITFMEELREGVALAARLRYCTNP
TVVPLNFTGHLGPIYRRIPEOGGMHVLITROXYEPNMOQYIFENVREDEPE
ATVVLINVDNAPITOMPERCDIPREGGTECTCKKYSVADAGEISTPMTOITS
ATNDEEYFELVRENIGOMKTVHRLINKLADLEENLHLFRTALDSISTPMTOITS
MVQENIESRPMMEIFAVQDFEKTAAQAEFVADIDGTDIDPIRIETESKD
LESVETIGAGREGAMFKAPIDRDITLEVEFVRLIAYKYGDNDVDESPSESTDIY
IIVNDVDOAPVPRPSYIEIMEEAMTLELDFEGFDRDLGHAQYTVLIESISA
GAHEAPYLAPEVGOROSFIETGONHMLDEVEFOKIORAVAIMDDPRVGLI
INILINMNDLPIFEHNVCTEVEAGFATVATKARDIIDDVESHLMGNVAV
YLSIDKTDGDLVITIDAFNHRONELFVQIRADDTIGEPNTNTAQVLQDINNT
PPLRLPRTTPSVEENPDEGVIPTELHADPDTTAEIRFSDMTSTATQGRDADA
EEFVNCIEIEFVNLNDRGTAIGRVVREIREHVTIDYEFVLVLTAVTDLNTVI
GDYDISEFTIIEIDMNDPPLWEGTLOEFREVREVAASGVIGSVLTALIDGLN
OVRTIIPRLTPELDIVEINDSGOISKKRQALDADDPPOHLXYTVVADKVDLLS
VDVCPDPNPNYNTGTDITIHITDNNPNPREDEKPEFVIVYVGADEGVNOLFS
DLNDEYTHKVSQINAINRPLRDOFEVDELEGLVYVYNNAGKLRDGDDEPHRF
FNVIDNEFGSDGNRNODEVLLVLDININPELPLGSDMISGLDQVRYTPI
FAPDRDEGDNSSVAVDIVLSPTDRDITLPQLFTITTEKDGIDQGTLEAFMDL
RGWATYIEHKAVDHPORISYEKPLVLRFPDVPFVPOPGMTILAEKAV
VNGVLTAYDGEFLRIVATDEGDLHAGVFSISGDDEALQYFVFNDSVGLALTT
QLPREDPREPVYTRANDGCEPERPSIDCTITVYVPTGCEPPEPSTYVAIERD
AGMERATLPLAKPDRNTMCDCHDYISTVIGSNKGFVADQSNLEFLTLERA
BOETHRTLIGASDPSPAVYLQASTLYTVVAVRANRPVQSLAYAGISTDTINR
GLTLNHTASGLPVTYTLVDSDMEADSTLQAVDEAFNLNPOGVLTANFQPSNH
GMFEEDVWATDTGTEARTKVKYVLISRNVEFTPMNLTEVEPNDDELTAEPTLF
GMRNINIOALPASPAPNAGAADQTEVRAHRIIDDLVPAREITOLGPNPITATION
ALQENLNLADLTFGETPIITGGEQAKRAVYALVAAALALCYVLLITLFFTRIRAN
RLRLALSMTKYSQSDGILNRVGLAAPGINKRAVAGSPINNETLKADFPALQSQSD
SGLIGIEDLPQFRNDYFPPEDESSMRGVYVNHMGANSVANHHNNFENATPFSPERA
NSOLRR"

BASE COUNT 1393 a 1433 c 1417 g 1112 t
ORIGIN
Query Match 35.4%, Score 1946; DB 3; Length 5355;

Db 597 CCAGGTATGACGACATCTCTTCACGCGTAGAGGTAGAGCATGAACCCGTACTGTGGCCAG 656
Qy 636 gtggagctgagatagatgaacatcgacagcaatccggccatctcaagtgctcagcgcc 695
Db 657 GTGATGCTCTCTATTTGTCAACATCGATGACAAAGATCCATATATACAGATGTTGAGCCT 716
Qy 696 tgcgaatccagaacatlgcgaagctagactgcgaactgcgtgtacaaagtgtcagac 755
Db 717 TGTGATTTCTTGAAGCGGCTGAACAGCATCATCATGCAAGTACACCGGAGCCAT 776
Qy 756 cgcgaagctgtaataagcaaccgcttcatagaagttccgtgtcgaagcagcagcgctgca 815
Db 777 GCTGACGGGAGATCAGTACACGTTTCATGAGGTTTCT--GAATATCAGCGCATGAGAC 833
Qy 816 gatgaagacatctctacatggtgttgagatacgaccccgagcagcgtgtcgaatagaag 875
Db 834 GATGACGAATATTTGCACTCGTCAGAGAAATATACAAAGCAATGATGATGTTCAAT 893
Qy 876 atgactgtggagatcaatctcgcccttgaaactgcgaacacactcagcttcataattagc 935
Db 894 ATGAGAGTTTCAGTCAAAAACCTCTTGTATATGAGAAACCCGCTACATTTGTTTACA 953
Qy 936 gtcaagctctgactgcgtaccgaacacacacaggtlccatgagtgatgcaagtgag 995
Db 954 GTTACAGCTTATGATTCCTTACCAAAACACATACAGTACGATGATGTTGCAAGTAGAG 1013
Qy 996 aacgtatagctcgcgcgcctcgctggtgtgagatctctcagtgagagatgttagcag 1055
Db 1014 AACGTGAACAAAGACCGCGCGATGATGAGATATTTGCTGTCAGATTCGATGATGAG 1073
Qy 1056 aagactaacacagactctccctccgcgcgcatagacgggagacaggaatcaataagagc 1115
Db 1074 AAGACGGAACATTCCTTTAGGGTTGAGGCATCGATGAGATACGGGATTCGATTAACCT 1133
Qy 1116 atcaactaacctcaatcgaagtgagcgtcagactcttct-----ccctg 1163
Db 1134 ATTTCTATAGATCGCAAACTGAAAAAGAGAGAGAACTTGTTCAGATTCAAAAGATA 1193
Qy 1164 gaggatgaaagcaggaactctctgcacgtgactgagatgcagccgacagcagcttga 1223
Db 1194 GAAGGTGTCGAGAAAGCCCTTGTTTAACGTCCCTCAATGACAGGACACTCTAGAG 1253
Qy 1224 agagagcttcaacctcaccatcgltgcttacaactactactgaagcgttagcttga--- 1280
Db 1254 AAGGAAGTTTCCACGCTGTCATTAATAGCGTAAATATGCGGATATGACGTGGAAGCG 1313
Qy 1281 -----acagagcccaactttcatcatcgtcaagcagcgaatgacag 1325
Db 1314 AGTTCGTATCCAGTCGAAACCGATGTGTCATCATCGTGAACGATGTCAATGATCAG 1373
Qy 1326 cgaaccgagcgcgtgcaataaagaatacatgatatcatgaggaagaactccaatgact 1385
Db 1374 GCGCCCTTCTCTTCGCGGAAGATCTCATTTGAAATATGAGAGAAACTGCGATGACC 1433
Qy 1386 ctaaatctcaatgaagaatttgatctcatalgacagatttgggtgaaacgctcaatc 1445
Db 1434 CTGAATTT---AGAAAGCTTTGGGTTCCATGATAGAGATTCCTGCTCAGCGACAATAT 1490
Qy 1446 acagtggaacttgaagagcgtgtcccgcaagggcggtccgcatcttcaatcgccg 1505
Db 1491 ACAGTACACTTATGAGAGCATTCATCTCCCGAGCTCACAGAGCGTTCATCAATAGACCG 1550
Qy 1506 gggagcgctacacagagagagacacttcacatgggacacataaacaacacagcgtgag 1565
Db 1551 GAGGTGCTTACCAAGCGCATGCTCTTATTTAGGACACCGACAGACCATCATGCTGAGAC 1610
Qy 1566 tacgaagatgcatcttctcaagaacatcatcatgaagtcgaagcagtgagacatgaaacac 1625
Db 1611 TTCAGAGTCCACAGATTTCCAGAAATATACAACTGAGGGCGGTAGCGATGACATGAGCAT 1670
Qy 1626 gctagcgaagtgaggcagcgctggtgtacgtgaacctgaactcagcggagcgaact 1685

Db 1671 CCCAAATGGGTGGTATGCGCATTAATCAACATTTAACTGATCACTGGAACCATGACGTG 1730
Qy 1686 cccatcttcgaagagagcagactcgcgtcgttcaagagacccgtcgcgccgctc 1745
Db 1731 CCGATTTGAGAGAGACGTGCAAACTGTACAGTTCGATGAGACAGAGGCGCCAGGCTTC 1790
Qy 1746 ccggtggccagagtgctcgccctcgacagagacacacagcagctagtagtcatcatg 1805
Db 1791 TATGTGCGCCACTGTGTGTGCGAAGACCGGATGTTGGATTAAGTCAACACTCTCA 1850
Qy 1806 atgggcaacgctggtgacactcgttcatagatgaatacagggagagatctcttgagc 1865
Db 1851 ATGGATACGCGAGTAACTGATCGATGAGATCGACAAAGAAACCGCGAGATATTTCTGACA 1910
Qy 1866 atgacatgctctcgaactacacacagacagaaactctatctgttcaggtgctgcgtgac 1925
Db 1911 GAAACGAAGCATTCACATATCAACAGGACGAAGCAACTCTTGTGAGATACAGCTGAC 1970
Qy 1926 gatacttggcgagcgccacacacacagtgacacacagcctgtgtatagaactgag 1985
Db 1971 GACACGCTGGGCGA--GCTTACAAACACCAACACTACTCACTGTGATCAAGCTGCGG 2027
Qy 1986 gatgtcaaacactctctccacccctcagcgttgcgcctgtgactccaagcgtgagag 2045
Db 2028 GACATTTAACAAACCCCTCTACGCTCAGGCTGCTCGCGCATCTCATCTAGTGAAAG 2087
Qy 2046 aacgttcccgagatagatatacccggaatctacgtlccacaccccgagacacagc 2105
Db 2088 AACGTGCCGACGAGGTTTGTGATCCCGACGAGCTGCACGCGACGCGCCGCACTACA 2147
Qy 2106 gctactctggttcgagatcgactgagactccacacttgggccaacgaagcgagagag 2165
Db 2148 GCTGACCTCGCTTCAGATTCAGATCGACTGCGAAGCTCTATGCTACCAAGCGAGCGAAT 2207
Qy 2166 accaacctcaactgaatcagtcggtgtatagttatagaaagatataccacccagagagc 2225
Db 2208 ACTGACTTAAGAGATATATCGGTTGTATAGAAATGAAACGATATACCGATATTAAC 2267
Qy 2226 aacccgggtctccgcacatcgcgccctcggtgtgcaagagatccggaacagctccacc 2285
Db 2268 CAGCGAGGACAGCCATCGCGCGCGGTGATGAGAGATCCGGGACGGCGCACATTA 2327
Qy 2286 gacttcgaggaatcgaatgtcttcaactcaccgctccgcgtgaaggaactcaactgtc 2345
Db 2328 GACTATGAGATGTTTGAATTCATCTACCGCTCATTTGATGAGGATCTCAACACCGTT 2387
Qy 2346 atcgaagatgactagatgagcagcgttcacgatcaacaataatcgaacatgaacaaac 2405
Db 2388 ATGGAGAAACCATATATATTCACATTTCAGATTCACATTAATGACATGAAACGCAAC 2447
Qy 2406 gcgcgactctcgcaagcagcagctgacgacagatgctgcgtgcaagcttgcgcgcgc 2465
Db 2448 CCGCCCTGTGGTGGGAAGCAACCGTCGACGCAAGATTCGCGTGGGAGAGGGGCGAC 2507
Qy 2466 agcgcagcgtcactcgcgtcgtgtcgcacccgacacacagcgcgcgtctcaaacaca 2525
Db 2508 TCAGAGATTGTTATAGATCCGACTGCGCATGATGAGAGGACGCGCTGATATATTA 2567
Qy 2526 gtgcgtacacataaacaactagaacaacactcccgaggatagtgagaattgac 2585
Db 2568 GTGCGGTATATCTATCTCCACACTAGACACTCCAGAAAGACCTAGTGACATAGACTTC 2627
Qy 2586 acaaactgtcaaatlgagtgagtgcaacgagcagatgcatgacagcaaaccttgcgc 2645
Db 2628 AACACGGGTACAGATTCGCTAAAGTTACACAGGCTATAGACGGGACGAGCGCGGCT 2687
Qy 2646 ttctactgtactacaacgctcagtcagcagagtgctcccttggaaaacgcgaagaa 2705
Db 2688 CAGAACCTTACTACACCGCTCATAGCTAGTGACAACTGTGACTCTCTTACTGTACTGAG 2747
Qy 2706 tctctcagatccaactcaactcgaagttcgaagtgatgataatagaataatcagac 2765
Db 2748 TGTCCGCTGACCTTACTTATGAGACACCGGAGAGATTACATCCACATTAACGAC 2807

QY	1755	tgctgcgccctcgacaagagacaatcgagaagctagtagtgcatactatgtatggtgcaacgctg	1818
DB	1781	tcagaggcttcacagaccggacacatagacacacgacctggcattttccattcttggamaatgcga	1840
QY	1819	ttgactacctgttcataigatgataatcaacccggagagatcttcgtgagcatgagacatgct	1878
DB	1841	atcacatttttgagaaatcacacamaaaactggcacatatacgtattatgtatgtacggct	1900
QY	1879	tcgacatcacaccgaagaacaccttatcttgaggctgcgctgtcgaatactatgtggcg	1938
DB	1901	tcgatttatcacagacagaaatgaatttaactacaaatgctgcgctgcagacacacatgtcgg	1966
QY	1939	a-----cggcccaacaacaacagltgacaacccagctggtgtagaactctgagagatgcnaa	1995
DB	1961	agccagactccacagccatpatacagccgctgctcagctgctatagaaactcgagagcgtcaaca	2020
QY	1996	aaacccctcccaaccttcagcttgcccccgttcgcgtccaagcttgagagagaaactcccg	2055
DB	2021	acacaccttctactacttgaggctgacctgacctgaagctcgtctgtagaagaaatgtgtccag	2080
QY	2056	aaggtatagagataatcccgaaatactctgataccgaccgagacacaaacgctactcgt	2115
DB	2081	agggctttgaatatcaacccggagataacgccacacagaccctgacacacagatattcctgc	2140
QY	2116	ggttcgagatctgacttggagctccaactggtgccaacgaagagggcagagagacaaacct	2175
DB	2141	agtttgaattagatttggagacacatcttcttcactaaacagggcgctgattaccatttccaa	2200
QY	2176	ctgaaatagctgggtgtatagtatatcgaaacgtataaccccaacgagaggaacccgggtt	2235
DB	2201	tagaacttccacggatggcgatataataaaacatcttccaaacccagccgacacacgaag	2260
QY	2236	cgcacatcgggcgctcgtgtgtgcaagataatccggaaacagctaacatctgactcgag	2295
DB	2261	agcgctgtggggcagatgctgacacaaaggagatccgccatattgtagacattcttggag	2320
QY	2296	aatcagagatgcttacctcaacgctccgctgtagaggacctcaactgctatcagagata	2355
DB	2321	agtttgaatttctctactaccatacagtgagattcgggactttgcacacagatagacgacag	2380
QY	2356	actacgatgagcgagcttcaacgtacacataatcgtacatgaaagacaacgcgcacatc	2415
DB	2381	attatagagaattctacaccttccacggtatataatataatattgaaacacaaactggcctatct	2440
QY	2416	tcggcagcgagcgtctgacgacagatagtcgctgtagcgagcttgcgcgacggcgagc	2475
DB	2441	ggcgctctgtgtttccttaaccacgaaccttccatgattttggagaggaattantttacccgctgc	2500
QY	2476	tcactgagctccgtgtctgcgcaacgacaatcgacggtccgcgtctacaaacaagtgcctaca	2535
DB	2501	tcattcgggctcgactacgtcgctacagacattgatggccactttacaaacaaagtcggctaca	2560
QY	2536	ctatacaacctagaacaacaaccccgaggagatagtgaagtgtacttcaacaatgtc	2595
DB	2561	ccatttatccccacaggaatattcttcgaaagctctatctccagataatttctgtttacgggctc	2620
QY	2596	aaattgaggtgagatgcgaacgagcgatcagatgcagaacgaaccccggtcctactatgt	2655
DB	2621	aaattttagcttgatgagaaatggtgcataatgcacgcctgatatttccacctcgtttggcacctca	2680
QY	2656	actacaacgcatcgtctagcgagctgtccctctggaaaacgcgacggaatgtcctccag	2715
DB	2681	actaacacgcttatagccagacaggaatgttctgaagaaatgaagaaacgaactgtccccgg	2740
QY	2716	attccacatcttcgaagtctccagcgagatcggaaatagaaatcatcgacaacaaca	2775
DB	2741	attccagctttcttggaataactctcgggacacatattatatttaccattctgacacattaaacaca	2800
QY	2776	aagtgctcgagcgcctcaactgagagaagtctaaacacggcgctgtgactcgtggagaatgc	2835
DB	2801	aggtcccgacagcagacactctagtcgattcaacgaacacggctgtacattttatgaaatgtac	2860
QY	2836	cgaagcgagacgaggtgtgtccagctgtactcccaacgcgttacaagagacgattgtacc	2895

Db	2861	CCGATTTCACGACACGTGGTCAAGATATATCCATCCAGCAAGACAGAGCAAGAAATATATAC	2920
Qy	2896	acaggtacacgtatacagctagaacttcgcgtgaaacccccacgcagcttccttcagag	2955
Db	2921	ACACGGTGGGTACCAAGATCATATTATGCTGTGAAACACACGGCTGCGAGACTTCTTGCCA	2988
Qy	2956	tgagacttgacacactgctgcctttagtgatatacccgggggagcagaanaattgagccgcg	3011
Db	2981	TGAGACTCGAATTCAAGGGCAGGCTGTAACGTGGAGAAACCAACAAATGAGCTCTCGATCGGG	3044
Qy	3016	atggagatgagcctacacatactacttglylaaattacatcgataactcttctcatg	3077
Db	3041	ACAGAGGGGAGAGACCACACACGGATATTCATTAACTCATTTGACCACTTTATAGCGAAG	3100
Qy	3076	gtgagcgtgaggaagacccaagagcaagtgtaaattatgctcgtctatgagtgagag	3133
Db	3101	GGAATGGAATATGAATATTAACACTCTACAGGGTGGCTGGATATCTTATATGATGAGAAATG	3166
Qy	3136	acaagcctccttgagatgcatctgcttgatgaactcgcgtttagatgttcccgaaagacag	3199
Db	3161	ACAACGCTCTCTGAATTTGGCGACTCCAGAGAGCTGGAATTTGGACCATTTCCGAGATTTAC	3222
Qy	3196	ttgctgtgtctc-----gttactccagaatctacgaccgagacagatgtaac	3244
Db	3221	AAGAGGATATTAACACTCGATGGCCAAAGACATGTGATATACGACCGGATATAGCAAAAG	3280
Qy	3247	cagaaacgaaacactcgcgtgtgtggttagaattcctgagccctcaagatcacagaccgcg	3300
Db	3281	AGGACACGCCAAACTCTCACCTTGGCTACGCCAATCTCGGCATGACAGTCCACAAATAGAG	3344
Qy	3307	acatcga---ggtgcgcgattctctcaactatctcgatctgtaaaacaaactggggaac	3365
Db	3341	ACCTGGAACACTGTTCCGAGACTTCTCAACATGCTGTGCGCTAACCAACGTAACCGGATTC	3400
Qy	3364	ttgagaccgtatgagacttgaagggtatgtgggacatttgcgaataatcatatgagcct	3423
Db	3401	TCCAGACAGCAATGCTTTGAGAGGATATTGGGGAGCTTACGATATATAGTAACTGGCT	3460
Qy	3424	tcgaccaagcgtaccgcgacgagaggtcccaagagactacacctgctacccgcct	3483
Db	3461	TGCACACAGGATATTCCTCAGACGATATCTCATGAGGATATGAATTTGAAATTTGCACTT	3520
Qy	3484	acaacttcacaccacctggtgtgtgtgtctccgaaaccgcctcgcctatcgcttcta	3544
Db	3521	ACAATTACAACTCTCCCTCAGTGTCTTTTCTTGAAATCCGGGACGATTTCAACACTGCTT	3588
Qy	3544	gggagcgcgcaacagaaagcgcgcttcctggcagcgcgtgcgcaacgagctcttggagccga	3603
Db	3581	TGGAAACGGCGAGTGTAAATTAATGTTTGTACACTTGTGAAACGGGTGACCCGTTAGACAGA	3640
Qy	3604	tctaagccacagaaaggaagggcctccagcgggcagcgtcactccgcctccaggaa	3665
Db	3641	TACAAGCAATTGACGACGATGCTGTTATCTCGCGTGGACTTTTCGATATTTGTTGGAG	3700
Qy	3664	atgaggaagccgttcagctacttgatatactgaatgagtggaagcaggaagaatagcgggc	3722
Db	3701	ATGCGATAGCTCAACTACTTCTGAGATTAATATATGCGACAGCTTTTGAACCTTGT	3766
Qy	3724	agctataattagccagcgtttcccgaggaanaatcagacaattcagatcacgactcggg	3785
Db	3761	TGCTGACA-----CAGGCGCTTCTGAGAAAGCAAGGAATTTGAGTTTACATCCGGG	3811
Qy	3784	ccaaggaagcgcgacaggaagcccgccgcctttggacgcagcgtcacgcttccgtgtct	3844
Db	3815	CTACAGAGCGGGGAACAGAACTTCGATCATATTTCAACAGACTTCCACTATATACAGTCTCT	3874
Qy	3844	tcgtatacccaacagagggagccacaggtgttcaacgaaatgacgtactgtgcctcttcg	3903
Db	3875	TGCTTCCGACTTTTGGGTGATCCGATCTTTTAAAGATTAACACTTACTCAGTAGCATTTGGT	3934
Qy	3904	aggggtgaagaagcctcgtgtagagttttagcgtgcgcaagcagaagcctttaaacc	3963

Db	33935	AAAAAGAGGTTGGCTTGACGTAGAGAGGTTCCGCTCCCAATGCAGAGAACCTTAGAACA	3994
OY	3364	acctctgcgaagatgactgcgaagatactactaagttatgacggcaacaagag	4023
Db	3395	AACCTCTGACCTGAGAGACTCTGCACGAATTTACTACAGGATCTTTGGTGTGTGATACG	4054
OY	4024	gtcttctgacttgagaccagtaagaagctactcccttgcgaaaggttgagccgcg	4083
Db	4055	AGCCATTGACCTTGACCTGGACCCGGTACGACAGCTGATCTCTGAAATCAGACTAGACCGGG	4114
OY	4084	agtgagcaagcttcttacaagcttgcacatcgcgcgcgaacactgcgcgcgcacttggga	4143
Db	4115	AGACACTGCGTAGCGCATGTGGTGCAGTGGCAGCCAGTAAATTCGCCACAGAGAGCGGA	4174
OY	4144	tccctctgcgaacttccactcgttgttaacggttcaatgtaagaagaacgcgcgc	4203
Db	4175	TACACACCTCCGGGTCTCTTCCACGCTACCTGTCACTGTACAGAGAGGGATTCACAGCG	4234
OY	4204	caattctgcgaagaccttctacacagcgagcaatttcgcgcgttggacagcatctggcgg	4266
Db	4235	CTGTGTTGAGACACCGCTGTACACGGCTGCGATTTCCTCCGATTAACATCAACAGG	4294
OY	4264	aattgtactgtltaaggcgagccacacagaagacgacacacatcagtaacataagacc	4322
Db	4295	AACATCACCGCTTCGCGAACATTCATTCGAAAGCCAAATTGACATATACATCAGAG	4354
OY	4324	gtcgcgaatgacgtctgagacagcactagaagcgttgcgcgcactgcgcgtcgcgtgc	4383
Db	4355	ACGGTTGATATGGCGGTGAGCTCCACTCTGMAACCCCTCAAGGACCTGGGGTTCCATCTGA	4414
OY	4384	atgagcacacccgggtgcttgcgttaataatgagccacgcgttccatagacagcatgt	4443
Db	4415	ACGGCGACACCGGGCTCTTATATCTGAGATACACCTTCTGCCACATATCAGGCGCATGT	4474
OY	4444	tcgagctcgacgtlcatcgctacgatacagactctgcaatcgacacagccgttgaaag	4503
Db	4475	TTGAGTTCAAGTATGATGCTATGACACCCAGATGAGAAAGACATACGGCAGAGGTGAAG	4534
OY	4504	ctctactcatctcatgcgaaaacgcgtgacttcaatttcgataccaacttgaagccg	4566
Db	4535	TCTACTCATATTCTATCCCAAAATAGGGTGTCTTCTTATCTCGAAGCATGTGGAGAGCG	4594
OY	4564	ttgagcagaagaagaattctaatagcgccacgttgaagcccggttcaactgaagctga	4622
Db	4595	TTGAGAGTACACAGACCTTATATGCGAANAAGTTACGCTTGACTTCAACATATGACTGCA	4654
OY	4624	acatcgaccagtgatgtgcgtlctcagcgacacgacgcgcgttgcgcgaagacgacacacgc	4683
Db	4655	AATATGATACAGGTCTGCTCGGGGCCACACAGACCGCGGGGTGATTCAGAGAGCCATGGGG	4714
OY	4684	agtgtcgcgcacttcatccgcggacaacgtgcgcgttgcaggaacaagaagtcgaagccg	4743
Db	4715	AAGTCAATGCTCACTTCACTATACAGGAATTAACATCCCTGTGACGCGCGCAATATTGAAGAGC	4774
OY	4744	tcgcgaacgaaacggtgcctctctgcgaaccaatccagctgtgtgttggacaccaacgcctgg	4803
Db	4775	TTTCGAGTAGACACTCACTCTCTGCGCTCGCTCAAGGTGTGTTGAACCAACGGCTGTTGG	4834
OY	4804	tgctcgaaagacttgtagacgggtgacacttcgacgcgttagcgagatgtaacgtgacatcg	4863
Db	4835	TTCGTAAGACAGACTGTGACGGGGGTCAAGCCCTGATATTCGGCACTCGCGCGTGCAGATCA	4894
OY	4864	ccgtctacgaactagccgcgcctctcgcgttgtgttagcttctcttgctcgtactgcttc	4923
Db	4895	CCATCTATGTGCTATGCGCGGGTGTGACACCACTCTTGCCTTCCGTGTGCTTATTTGTCGCA	4954
OY	4924	tcgcatgttctttagaacaagagacactgaacccgcgcagctctgaagcaatcttcatgacga	4983
Db	4955	TCACATTCATATGTGAGAGACCCGAGACTCTGTAACCGCGGTTGGAAAGACTGTGATGACGA	5014
OY	4984	agtagcgtcgtgtgactccggcttgaaacgcgcgcgcgtt---ggcgcgcgcacacaca	5040
Db	5015	AATACGGCTCGTGATTTCCGGGCTGAACCGATGGGGGTATACGGCGCCACGAGAACCAACA	5074

QY	5041	agacgcgcgtcgaagggcccaaaccccatgtagaacgagcgccctcgcgcgcgcacttcg	5100
Db	5075	AACACGCCATCGAAGAGGCTCCAAACCCATCTGGAAAGAGAGATCAAGGCCCGGACTTGG	5134
QY	5101	aaccatcatgtagcgcgagtgtagcactccagacctgatcgacatcagagatgcgcga	5158
Db	5135	ATGCCATCATGTGACACATCTGACGACTCTGATCTATTCGGCATTCAGGATATGACTGCA	5192
RESULT 12			
	AX147207	AX147207 1604 bp DNA	PAT 08-JUN-2001
LOCUS		Sequence 7 from Patent WO0136639.	
DEFINITION		AX147207	
ACCESSION		AX147207.1 GI:14346378	
VERSION			
KEYWORDS			
SOURCE		European corn borer.	
ORGANISM		Ostrinia nubilalis	
		Eurayota; Nubialis; Arthropoda; Tracheata; Hexapoda; Insecta;	
		Pterygota; Neoptera; Lepidoptera; Lepidoptera; Lepidoptera; Glossata;	
		Ditrysia; Pyraloidea; Pyralidae; Pyraustidae; Ostrinia.	
REFERENCE		1 (bases 1 to 1604)	
AUTHORS		Flanagan,R.D., Mathis,J.P. and Meyer,T.E.	
TITLE		Bt toxin receptors from lepidopteran insects and methods of use	
JOURNAL		Patent: WO 0136639-A 7 25-MAY-2001;	
		PIONEER HI-BRED INTERNATIONAL, INC. (US)	
FEATURES		Location/Qualifiers	
source		1..1604	
		/organism="Ostrinia nubilalis"	
		/db_xref="taxon:29057"	
BASE COUNT		386 a 465 c 419 g 334 t	
ORIGIN			
Query Match 28.1%; Score 1547.4; DB 6; Length 1604;			
Best Local Similarity 98.1%; Pred. No. 1.3e-262;			
Matches 1566; Conservative 0; Mismatches 31; Indels 0; Gaps 0;			
QY	3040	tcttgtaatttcgcataaactcttcttcgatgctgtagagtagaagaacagagacg	3099
Db	8	TCTTTTCAACCTCATCGACAACTTCTTTCGAGGGTGCAGGTAGAGAAACAGAGACG	67
QY	3100	aagtgaaatattgctgcttcattatgtagatgtagaacaacgctccctagatgcatgac	3159
Db	68	AAGTGAATATATTTGCTGCTTATGTGATGTGAAGACACACCTCTCGATGACATCGC	127
QY	3160	ctgtagaactccggtttgatgtttcccgaaagagcggtgtgtgtgttcgtgtaactccag	3219
Db	128	CTGAGCAACTCCGGTTTATGTTTCCGAAGGAGCAGTTGCTGCTGCTGATCTCCAG	187
QY	3220	aaatctacgcaacgcgacaggggagtagaacgagaacgacaactcgcgtgttaccgaa	3279
Db	188	AAATTTAGCGACCTTACACAGGATGAACACGACGACACTCGGCTGCGTTACGGAA	247
QY	3280	tcttgacccttaacgtaacgcgacgcgagacatcgaggtgcggaactcttaccatgatct	3339
Db	248	TCTTGACCTTCACGATCACCGACGACGACATGAGAGTGGCGGATCTTCCACATGATCT	307
QY	3340	cgattgaaacaaactcgtgggaacttgagaccgcgatgacttgaggggtatttgggga	3399
Db	308	CGATTGAAACAAACCTGGGAACTTTGAGACCGCTATGACTTTGAGGGGGATTTGGGCA	367
QY	3400	cttaagaataattcatagagccttcgagacacggttaccgcagagagaggtccaagaga	3459
Db	368	CTTAGCAATATTTCTATTAGAGCCTTTCGACACACGCTATCCCGACGACAGAGTCCAAAGGGA	427
QY	3460	cgtacacctgtgtcatcgcgcctctacaacttccacacacctgtgtgtgttccgcgaac	3519
Db	428	CGTACACACTGGTCAATTGCGCCCTCAACCTTCCACACACCTGTGTTGCTTCCGGAAC	487
QY	3520	ccgactccgtaactcgcgttctttagagagcgcgcaacagagagcggtctctcgcgaagc	3579

D	488	CCGACTCCGTCATTCGGGCTCTCTAAGSAGCGGCCAACACAGAGCGGGGCTCTGGGACAGG	547
Q	3580	ctgcacaaagatctccctggagccgactctacgcacacgcagagagcggctctccacgcggcgca	3639
D	548	CTGCCACGAGTCTCTCGAGCCCGCATCTACGCCACCGACAGGAGCGGCTCTCACGGGGCA	607
Q	3640	gcgtcaagcttcacgtcccaagggaatgagagcgcttaagtaactttgataaactgaag	3699
D	608	CGGTACGCTTCACGCTCCACGAGGAAATGAGGAGCGCTTCATGTAATTTGATATTAACGTGAG	667
Q	3700	tggagcagagaaatacagcggcagcttatactaacgcagctttcccaagcgaatca	3759
D	668	TGGGACGACGAGGAAATATAGCGGGCAGCTTATATACGCCAGCTTTCCAGACGAATATCA	727
Q	3760	gacaattcaagatcatcgatcccgycacagcagcggcgycacagccggccggctttgta	3819
D	728	GACAATTTCAGATCAGCATTCGCGGGCCACAGACGGCGGACAGGAGCGCGGCTTTGGA	787
Q	3820	ccgacgtcagtttcgtggtgtcttcgtlaccacacagggcgaccagtgltcagcgaa	3879
D	788	CCGAGCTCACGCTTTTCGGTGGTCTTGTTACCCACGACGGGCGACCCAGTTCACAGCGAA	847
Q	3880	atgcaagctacgtctgccttttttggagggtgagaagcgctccgttgtaggttttgctgc	3939
D	848	ATGCAAGCTACTGTTGCTCTTCTGAGGGGTAAACAAAGCCCTCATGAGATTTTGGATGAC	907
Q	3940	cgcagcagaaagaccttaaaaacacactctgcgaatctgacgcagaatattctactca	3999
D	908	CGCAAGCAGAAAGACCTTAAAAACCACTCTCGCAAAATATCTGCCAAGTATCTTCTACA	967
Q	4000	ggttattgacgcgaacaacagaggtcttctgctacgtgacagcttaagcgaagctac	4059
D	968	GGTTATTATGACGGCAACAAAGAGGCTGTGTTGCTGTGACCAAGTCAGCAACGTCATCT	1027
Q	4060	cccttgcgcagggatcttgaccgcggaggtggcgcgctctaacgcgtgcacatcgcgcga	4119
D	1028	CCCTTGGCGAGGAATTTGACCGGAGGTTGCCACGCTTACACGGCTGACATTCGGGGGA	1087
Q	4120	gcaactcgcgcagcgcacacggggtaccctctgcagacttcaactctgtttgcacgtca	4179
D	1088	GCAACTGGCCGAGCGCACACGTGGGATTCCTCTGAGACTTCATCTCTGTTGCAAGGTCA	1147
Q	4180	atgtaagagaagcgaaacgcgcaccaatttcgagcaggaaccttacaacgcggacatt	4239
D	1148	ATGTAAAGAGAAGGCAACCCCGCCCAATTTTGAGACGAGGACCTTATACAGCGGCAATTT	1207
Q	4240	cgaagcttgagcagcatctggccgggaatttgctactacgtcaagcgcggccacacagaagc	4299
D	1208	CGAGGTTGGACAGCATTTGGCGGGGAATTTCTTACCTGCACGGGCGAGCCACACAGAAAGCG	1267
Q	4300	acacatcaacgtatacacatcagacgctgycagatgtagagtgaacagcagccttagaagc	4359
D	1268	ACACATCAGGTATCATATGACCGGTGCGAGCATGTAGCTGACACGACCTTAGAAGCG	1327
Q	4360	tggcgacatcgcccttcgcgtctgcatgacacacgcggcgtgcttcgtctcaatgcaac	4419
D	1328	TGCGCGACTCGGCTTCACCTCATGCTCATGCAACACCGGCGCTTGCCTCAATATGACGC	1387
Q	4420	ccacagcttcatacgaacgagatgtctgaattcgaaagtcatactgctcggataacagctctg	4479
D	1388	CCACCGCTTCATGACAGGAGATTTGGAATTTGAGCTCATTCGCTATCGGATATACACATCTG	1447
Q	4480	caatcgacacagcccggtgtaagcttactcatctcatctcgcaaaacgcggctgaccttca	4539
D	1448	CAATCGACACAGCTCGGTGAAGTGTACTCTATCTCATGCAAAACCGCTGTCTTTCA	1507
Q	4540	ttttcgtaaaccaacttgagacgcttgagcagaacagaatttcaatagcggccacgttca	4599
D	1508	TTTTTCGATTAACCAACTGTGAGACCGTTGAGACAAACAGAAATTTTATGACGGCCAGTTCA	1567
Q	4600	gcaccgggttcaacatgaagctgtaacatatgcaccagct	4636
D	1568	GCACCGGGTTCACTATGACTGTGAATATGACCAAGT	1604

[illegible]

Dh 220 CACAGCAGATCGATGTGCTACATGTAGAAATACCAGACGACGCGCTGACTTCC 279
Qy 280 cgcctatgttactgaaggaanaaatcatgtgagtgagcccttaatacccgccgaccc 339
Dh 280 CACCTCAAAATTTTGAGCGTTTAACTGAGGCTCGACGCCACTTATACAGCTAGGATC 339
Qy 340 gagaagaatgtatgtatgagaactcttaaccgataaata-----ttc 384
Dh 340 GAGAAAGAGTCTGCGCTCAATGACTATGAACTGATCCCTGAGCAACAACCATGCTGACC 399
Qy 385 aggtcatatcatatgtgaggaagaatcgaagagacgctcatatgtcggaagtcatatc 444
Dh 400 AGAGAAATTTACATGAGAGAGATCGAAGGTCCGTAAGTCTGAGTGAAGAAATTAATAC 459
Qy 445 aaggtccaacacgcgcgtgtcgtatgtatgtcagcagccagccagagccca---gctg 501
Dh 460 AAGGAACACCCCTTCCTCAATTAAGATTACTTTCGTTGGTGACACCATGCTG 519
Qy 502 ggccttgatctcg---acagaaatgaagcagcgcaatgagccttgtaatacga 558
Dh 520 GAGCAGAAATTCGTAATATCTGACCACTGAGAGACTGATCTTGTAAATCTCAAA 579
Qy 559 gacaagctagagacagcaaccctgcagagctatgtgttcaatccaagtgaaggtg 618
Dh 580 GGCAGGACTATGAACCTCTGATATGAGAGATACGCTTGATGTGATGGAAGGCC 639
Qy 619 aatacagagccgtactgtgtgctgtgagatagtcacaatcagacacatccgccaatcc 678
Dh 640 AGTGGCTGTGTAAACGTGTAGGCTGATATTTGTAACATGAGACAAATTCGCCATCA 699
Qy 679 tgcgaagtgtcagcgctcgtaattccaagaacatgycgaggttagactagcgactcg 738
Dh 700 TTGAGATGTGTAGACCTTTCGAATACCGGAACCTGTTGAACCCCATTTCAAAATGTA 759
Qy 739 tgtacaagttcagacgcgcagcggtgaatcagacccgcttcatgtcgctgtcg 798
Dh 760 AATATATGTGTCCGACGACGAGCTCTGATCAGTACAGAGTTGATGATTATCATATAG 819
Qy 799 aacaagcagagcgctgcagatgaagacatcttcaatgtgtgagaaatacagaccacg 858
Dh 820 ACACGAGAG---AGGAGACGAAAAGATTTGCAACTGATCGAATAATATCCGGCG 876
Qy 859 acgtgttcaatataagaatgactgtgggataaattcgccctgaacttcgagacactc 918
Dh 877 AATTGAGCAAGAGTGTATGTGTTGTAATGAAAAAATCTTGATTAACGAAGAGATC 936
Qy 919 agcttcaatattagctcagacttctgactcgtcgcgaacaacacacgcggtacaca 978
Dh 937 CTCTACACATTTTCAGATCGACGCTTCTGATTCCTTACCAAAACAAATAGACCGGTCA 996
Qy 979 tgaigtgtcagttgagaagctagagctcgccccctcgctgggtgagatctctcag 1038
Dh 997 TGAATGTGAAGTAGAGAACTGGAAACATAGAAATCTCTCGGTGATGAGATCTTTGCTG 1056
Qy 1039 tgcagcagtttgcagagagaactaactaactcagagcttccctcgcgcgagtagagagaca 1098
Dh 1057 TGCAACAGATTGATGAAAAAAGCGGAATATGTTACAGTGCAGACTATGATGCGGACA 1116
Qy 1099 cgggaatcaatagagccatcaactataccctcatagagatgaagctgaagactcttt 1158
Dh 1117 CGGGAATCAATTAACCTTATTTCTATGAGAAATGGAAGATGAAGCAAAAGATCT 1176
Qy 1159 -----ccctgaggtgtatgagaagcagactatctgacgtgagtgatcg 1206
Dh 1177 TCAGCATTTGAGAACCTTAGGGGAAGCAGAGACAGAGCTGCAAGTCCAGCTGGCTCTATAG 1236
Qy 1207 accgcagcagcttgaagagagcttctcaactcacaactcgtg----- 1252
Dh 1237 ACAGAGACTTACTGAAAAGGATATGTTTCAATATGAAATATTCATATAAACAAGGTG 1296
Qy 1253 -----ttaaatctctgaagctgactgttcgaacagagcccaacttctcatcgtca 1308
Dh 1297 ATAAATGCAAAAGAGTGAATCATCTGATCGAGACCTCAGCAAAATGTGACATTAATTA 1356

Qy 1309 acgagctcaatgtatcagagaccgagccgctgtcataaagaataglatgtatcag 1368
Dh 1357 ACGATTAATATGATCAGAGCGCAAGAACCTTCCATTAAGAAATACAGATCTCCATATGG 1416
Qy 1369 aggaatctcaatgtacttaacttaatatgaagatgtgatatgaatcagacgatttg 1428
Dh 1417 AAGAAATGCGATGCTTGTGATTTG---CAAGATTTGGTTTCCATGACCGTGAATTTG 1473
Qy 1429 gtaaaacgctcaatatacagttgaacttgaagcgtgttcccgccagggcgcgctcg 1488
Dh 1474 GTCCCAAGCCTCACTAGACGCTTACCTTAGAGATATACAGCCAGAGGGGCCCATACCG 1533
Qy 1489 catctacatcgccgagggagcggtacacagagggagacttcaatgaatggcaataa 1548
Dh 1534 CTTCATCATGCGCCCTGAAGAAAGTTACAGCGCCGCTGTTCCACTGATGCTAGAA 1593
Qy 1549 acacacacatgtgtatgaagatgtcaatt---ttcgaataatcaatlaagtgca 1605
Dh 1594 TCCATTAACATGTTGGATTTAGATGACGACTACAGACGAGATTAAGCTTAAGGCGAG 1653
Qy 1606 aagcagttgacatgaacaacgctagccaagtgagcgagcgctgtgtatcgtgaactga 1665
Dh 1654 TAGCAATTTGACAGACAGATACATACATCATTTGGGGAAGCAATTAATTAACCTTA 1713
Qy 1666 tcaacttgaagcagcaacttccactcctcgaagagagactcgcgctgtttaa 1725
Dh 1714 TCAATTTGGAATGATGAGTACTATATTTGACGAGAGACGCTTACAGCTGACATTTAGG 1773
Qy 1726 agacgctgcgcgcggtctcccggtgagcgagcggtcgcgcctcagagagaactcag 1785
Dh 1774 AGACGCTGCGTGTAGGCTTCACATGATGTAATACCGGGCTTAAGACAGACATTCGGTG 1833
Qy 1786 acgtagtagtgcattcatatgtgagcaacgctgtgtactactgttcatatgaatca 1845
Dh 1834 ACATATGAGACACTCATATTTGGGCAACGCTGCAACTCTCTGAGATTAACATTAAGTA 1893
Qy 1846 cgggaagatcttctgtgagcagatgccttcgactaccacagagaactctat 1905
Dh 1894 CTGAGAGTGTAGCTGTCACAGGACGATTAATCTTGTGATTAACAAAGACAGAAATCA 1953
Qy 1906 ttgttcaggtgcgcgtgtgagacttcttggtgcagagcccaaaacagttgacccc 1965
Dh 1954 TAGTTCAGATTTCTGCTGTATACCTAGG---TTTACCTCAACACAGGCTACCAAC 2010
Qy 1966 agctgtgtatagaactgtgagatgtlcaaacacactcctccacctaactcgttcccg 2025
Dh 2011 AGCTCAGATATTTTGGAAACATCAACACAGCCACTATCTGAGACTGGCAGGTT 2070
Qy 2026 cgaactcaagcgtcgaagagaacttccgaagatacagatatcccggaataactgt 2085
Dh 2071 CCAGTCNAAGTGTAGAAAGAAAGCTTGAAGTGGGACACCCGATTACGAGGGGCTAACGG 2130
Qy 2086 ctacccagcccgagacacagcgcctactcgtgtgtcagatctgacttggagactccac 2145
Dh 2131 CGACAGAGCTCAGACACCAACGATTTTACCTTTCGAAATGATGATGGGCAATTTACG 2190
Qy 2146 ccaaccaagcaggy---cagaagaccaaaccctactgaatacgtcgtgtatgtatcg 2202
Dh 2191 CTACGAAACAGGACCAATAGACCAACACTGCAAGATACACAGGAGTGTGAAATACC 2250
Qy 2203 aaacgatataccccaacgaggaacggtgtccgcacatcggcgctcgtgtgcaag 2262
Dh 2251 TGACGGTATACCGAGATCTGACAAATCAACGGAGAGCTGAGAGGCTGCTGACGCTG 2310
Qy 2263 agatccggagacaaggtacacatcgactcgcaggaattcagatgttactacacgctc 2322
Dh 2311 AGCTAGATGATGGCTGACCATGATTAAGAAAGTTGAGATGCTGATCTGCTGCTCA 2370
Qy 2323 gctgtaggactcaaacctgtcatcggaatgtactacgaatgagcgagctgtcagatca 2382
Dh 2371 GGGTATGATGCGCAACACTGTCTATGGCCCTGATTAAGAGCAAGCAATGCTGACGGTGA 2430

[illegible]

Db	3511	TCCTGGATTTCATTGAGACCTACCACTAACCGCTACAGGCGCATACCACTTCATTCAACGG	3570
Qy	3502	tggttggtgtcccggaaccggaaccgcgttcgtggtttctaaagagcgcgaaacagaag	3561
Db	3571	tgtttgtgtgtcccaactcttgcgttcaaacctacagctttttaaaggagcgctgctatntrtca	3630
Qy	3562	gcgcgtctcttgcgagcgctctgcacac-----gattccctcgagccgatctacgcga	3612
Db	3631	ANGGATAGCTGGCTCTGGCTTAATATCGGAGGCGGAGAGTTCTCTCGACAGACTCTCGCA	3690
Qy	3613	ccgacgaggaagcgccctcccaagcgcgagcagtcacgttccagctccagggaaatgagag	3671
Db	3691	CTGATGAAGATGGGCTACACACGACGAGAGTAACCTTCTCCATATAGCTGGAAACGATGAAG	3750
Qy	3673	ccgttcagtaacttgatataacttgaatgagtgagacagagaaatagcggcgtctatat	3732
Db	3751	CTGGCGAATATTTTCAATGTGT-----TGAAACGACGCTAACAATCGAATGTCTACCG	3804
Qy	3733	tacgcagcagctttcccaagagaaatcagaatctcagatccagatccgagcgacgagcg	3792
Db	3805	TGAACGACGACTTGCCTCGCTGGCTCCAGACAGTTTAAAGTTGGTATTCGGGCGCACGGACG	3864
Qy	3793	gcgcgacgagagcccgcgccgctttgacgcgacgctcagcttttcggttggtcttcgtaccca	3852
Db	3865	GGGGACGAGACCGGGACCTAGAGAGTACGACACTGTCCGTCACCTGGTGGTTGTGATGA	3924
Qy	3853	cacagcgcgaccacagtggttaacgagaaatgacgtactgtcgctctctcgagagtgag	3912
Db	3925	CGCAGGGAGACCCCGTGTGGACGACACACACTCTCTCTCGCTGTGTGAAGAGAG	3984
Qy	3913	aaggcctcgcgagagtggttgagctgcgcgaagcagaagaaccttaaaacccctctgcg	3972
Db	3985	CTGGATGTCCGAAAAGTTTCACTGCTGCTTAGGCGCATATCCCAAAACTTACAGGTGA	4044
Qy	3973	aagatgactgcgaagatatctactaaagtttatcgcgcgaacaacagaggtctctcg	4032
Db	4045	TGGAGCACTGCCATACCATCTACTACTCTATGTTGATGGCAACATGATGGTACACTTGG	4104
Qy	4033	tactggaaccagtaacgaagcatctccctctgcgagagtgtagcccgcgagatgagcca	4092
Db	4105	CCGTGGACCCGAGACTTACGCTATCTATTGCTGAAGCGCGGTGACCGCAACCAACAG	4164
Qy	4093	gcctctaacgcctgcacatcgcgcgagcgaacatcgccgcgaacactggaatccctctgc	4152
Db	4165	AGCAGTACAGGGTGTGTGTGGCGCTTCCAAACCGCTTGGCGGAC---CTCCACCTTGT	4221
Qy	4153	agacttccatccctctgctcgtcacaagtcacgttaagaaagcgaaccgcgcgccaatcttcg	4212
Db	4222	CCCTCTACTCTCTACCGCTCACCATCGGGGTTGGAACCAACCAACCTTAGACGATCTTGG	4281
Qy	4213	agcaggaaccttatacagcgcgatcttcgaagtgtagcaagatgagccgggaattgctta	4272
Db	4282	AAAGTGAATTTTAAACAGCTGGCGCTTACACACGATAGCATACCAAGAGAGCTCGTTT	4341
Qy	4273	ctgcgcagggcgagccacaaagaaagaaacatcaagcaacaaatagaacacgctgcgcga	4332
Db	4312	ACCTGGCGGCAAAACATTTCGAAGGGCTTCTATGCTCTACTCTGATAGATCAAGAAACCA	4401
Qy	4333	tgcaactgcgaagcagaagcctcgaagccgctgcgactcgcgctcttcgcgtcagtcagcca	4392
Db	4402	TGAATAATGACAGAGTCTGTTCCAAACAGTTGTGAGAGAGCGCTTCACATTAACTCTGAA	4461
Qy	4333	ccgcgcgtgcttcgctcaataatgagaccacgcgttccatgatcgaagcaatgcttgagctcg	4452
Db	4462	CCGAGGACATATGCGCTGAACCTTCCACCAACATCTGTATGACGGCGAGTTTCCACTTTTG	4521
Qy	4453	acgcacatcgctacgataacagctcttcgaaatgaaacagcccggttgaaatctactcca	4512
Db	4522	AGGTGTGGCTTAGTGACACCGCTGGAGCGAGTATCGACAAATGTGCATTTACATGA	4581
Qy	4513	tctatcgcgaacacccgctgaaccttcaattcttgataaaccaacttgagacgcttgagcaga	4572

Db 4582 TATGACTCGGTTAGAGTAGCTTCTCTTCTACAAACGGAAGCTGAAGTTAAGAGA 4641
 Oy 4573 acagaattatcagcgcccggttcagaccgggttcaaatgacgtgcaacatcgacc 4652
 Db 4642 GAAAGAAATTATTCACAAACGTTCCGCAACGGCTTTGTATGACATGTAACATAGACA 4701
 Oy 4633 aggttggtccgttcagcgacgacgagcggtgagcgcaagacgacacacgaggtgcgcg 4692
 Db 4702 GCGTCTCCGCGGATCCGACGACCGCCGATTCGCGAGGGGTACACAACTCCGAG 4761
 Oy 4693 cgcactcaccgagcaacggtgcggtgcagcaagaggtgcggtccgcgacgcg 4752
 Db 4762 CTCACCTCATACGAGACGACGCGCGGTCCGACGCAATATTGAGGGATTATTACGG 4821
 Oy 4753 acacggtctctgcgcacccatccagctgacgtgacgacacacacgacctggtgcgcag 4812
 Db 4822 AACTCAATACATTTGGTGTGATCAGAGAGGTACTGATCAGCAATTTGACGCTACTGG 4881
 Oy 4813 accgtgtagcggtgacactccgacgacgagcgaggtcaatgacagatcgccgtcagc 4872
 Db 4882 ACTTGGCGGCGGAGGATCGGCGAGTCTGCCCGGAGAGATAGCCGCTAGCGGTGACA 4941
 Oy 4873 cactagccgctctcgcgtgctgtaggtctctctctcgtcgtactgcttccgcatgt 4932
 Db 4942 TCCGCGCGGATCGACGCGTACTCGCGGTACTGCTGCTCCTCCATCCGCTTCT 5001
 Oy 4933 tctgtagacaaagacactgacgacgagcgtgcgaacgacctccatgacgagtcgcgt 4992
 Db 5002 TCATTAGGAGAACGAACTGAAACCGCGCATCGAAGCCCTCAATCAAAAGATGTTCTTA 5061
 Oy 4993 cgtgtgactcgcggtgtagcgcgacgacgagcggtgcgcgacgaacgacgacgcgcg 5052
 Db 5062 CGGACATGAGCCAAACACGCGGTAGTAGCATGCTTAACATTAACAAAGCACACAGAAC 5121
 Oy 5053 agggctccaccccatgtgaaagagagccatccgcgcgcgcgcgcgcgcgcgcgcgcgc 5112
 Db 5122 CTGGTTCGAATCTCTTATTAACCGGATGTTAAGACACCTTAACCTATTAAGACG 5181
 Oy 5113 acgcaagagcgacccgacccgacgtgacgtgacgtgacgtgacgtgacgtgacgtgac 5172
 Db 5182 A-----AGTATCCCATGACCTGTTGATGTCGAAGACTTGGAAACGATT 5235
 Oy 5173 actccgc 5189
 Db 5236 ACTTCCACCCGAAAC 5252
 RESULT 14
 AF367363 7799 bp mRNA INV 12-AUG-2001
 Locus AF367363 Heliothis virescens truncated cadherin-like protein mRNA, complete
 cDNA
 AF367363
 AF367363.1 GI:15149241
 ORGANISM
 Heliothis virescens
 tobacco budworm.
 Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
 Ditrysia; Noctuoidea; Noctuidae; Heliothinae; Heliothis.
 REFERENCE
 AUTHORS
 TITLE
 Gahan, L.J., Gould, F. and Heckel, D.G.
 Identification of a Gene Associated with Bt Resistance in Heliothis
 virescens
 Science 293 (5531), 857-860 (2001)
 JOURNAL
 PUBMED
 11486086
 2 (bases 1 to 7799)
 REFERENCE
 AUTHORS
 TITLE
 Gahan, L.J., Gould, F. and Heckel, D.G.
 Direct Submission
 JOURNAL
 Submitted (31-MAR-2001) Dept. of Genetics, University of Melbourne,
 Parkville, Victoria 3052, Australia
 FEATURES
 source
 1. 7799

CDS
 /organism="Heliothis virescens"
 /db_xref="taxon:7102"
 /note="r1 allele"
 194..2062
 /note="truncated CalP"
 /codon_start=1
 /product="truncated cadherin-like protein"
 /protein_id="AK85199.1"
 /db_xref="GI:15149242"
 /translation="MAVDVRLITAAVLIILAAVLTAFADQDSYVAIPRPDPDPNPF
 EGVPMSONPLLPEDREDYCMNAPDSALNPVTIVTMEEDIEDVAIALNKGHTP
 TVTTPNFTGHLGPVIRIPGQGDMLVITQRODYETPNMQITLVNREDEFOE
 ATVMILVINIDNAPMIOMPECDIDHEGTTECKSVSDAGISIFMFEIES
 DRNDEEYFLRENIOGMVYHMRLLINKPLDYENPLHFRVALDSIPNVHTM
 MGVENIESRPMEIEFAVOQDEKTAOSFRVALIDGPDIDKPIFYRIETESKD
 LSEVENTIAGREAGMFKVAPIDBDTLEKFEVHSLIAYKGDNDVGSFESKNDIV
 IIVNDNDQAPVFPFRPSYFIEIMEETAMTINLDEBFHBDLDPAHQIYHLESHPA
 GAHEAFYIAPEVGYQROSFVGTQNHMHMLDFEVPEFOKIQLRVALIDMDPRVGLAI
 INININMDELPIFEHDVQTVFKEEGEGRVATVLAQRVIDRVEDHSLMGNAV
 YLSIDKDTGDLIVTIDAFYHQNELFVQIRADDTLGEPIYNTNACSLSHRAGL"
 1982..2007
 /note="r1"
 complement(2029..2052)
 primer_bind
 primer_bind
 /note="R3"
 2030..2281
 /note="LTrA"
 2030..4292
 /organism="Heliothis virescens"
 /db_xref="taxon:7102"
 /transposon="Hel-1 retrotransposon"
 /note="r1 allele"
 2221
 /note="polyA site for 2.1 kb transcript"
 4036..4292
 /note="LTrB"
 4235
 /note="polyA site for 4.4 kb transcript"
 complement(4322..4351)
 primer_bind
 /note="R2"
 2100 c 1921 g 1791 t
 BASE COUNT 1987 a 2100 c 1921 g 1791 t
 ORIGIN
 Query Match 24.8%; Score 1363.12; DB 3; Length 7799;
 Best Local Similarity 64.4%; Pred. No. 2.8e-230;
 Matches 2170; Conservative 0; Mismatches 1158; Indels 42; Gaps 7;
 Oy 1952 caccagtgacccaccagctgtgtagaactgagagatgtcaacaacacctctccaccct 2011
 Db 4288 CTCGAACACTCGCCCAACGTGTGATACACTGCAGACATCAACACACCTCTACGCT 4347
 Oy 2012 acgctgtcccggttcagatccaagctgagagagaaagctccggaagatagagatc 2071
 Db 4348 CAGACTGCCCGCACAGCCCGCTCAGTGAAGAGAACGTGGAGCGGCTTGATCC 4407
 Oy 2072 ccggaatactactgtacaccccgagacacgacgacgacgacgacgacgacgacgacg 2131
 Db 4408 CACCGAGCTGCACGCGCTCGACCCCGACACACCGCGAGCTTCACATGCACTG 4467
 Oy 2132 ggaactcactgagccacacgaagcagagagacacacacctactgaatcgtcgcgtg 2191
 Db 4468 GGCACACTCTCTCGCCGACCAAGCAGGAGGATCTATGCTAAGGAGTTGTAATTG 4527
 Oy 2192 tatagtattcagaagatatatacccccagagagcaacggggttcgcgcacatcggcgcct 2251
 Db 4528 TATAGAAATCGAGCGGTATACCCGAACCTTGACACCGAGGACCGCATCGCGCGCT 4587
 Oy 2252 cgtgtgcaagagatccgggacacgtacacactcgcacttgcaggaattcgagatgctta 2311
 Db 4588 GGTGGTTGCGAGATCCGGAAACAGTCACTATAGACTACGAGATGTCGGGCTCTGTA 4647
 Oy 2312 cctcagcgtcgcgtgtagacctcaacactgtcatcgagatgactcagatgagcgacg 2371

Db	4648	CCTCAGAGTCAAGGGTCAACGGATCTCAACACGGTCATTGGAGACGACTATGATATATCAAC	4707
Oy	2372	gttcacggtccaaataatcgacatgaaacgaacaaacgcgcgcatcttcgcgaacgacgct	2431
Db	4708	ATTGACATCATATATATATATAGACATGAACGACAAACCTCCGGCTGTGGGTGGAAGGACACTCT	4767
Oy	2432	gaacgacagatctgcgttcgcgaagcttgcgcgcgaacgacgacgtcactgcgtccgct	2491
Db	4768	GACCGAGGAGTTCGGCGTGGAGAGGTCGGCGCTCAGGAGTATATATATGATATGTCGGTACT	4827
Oy	2492	cgcaacggaacttgaacggccgcgtcttcaaaacgaafgctgtacactatcacaactcagaaga	2551
Db	4828	CGTACCGATATTCACGGGACCTCCTTATATATCAAGTGGCGGTATACCATATCACTCTAGACT	4887
Oy	2552	caacaactccgaaggatctagtgaaattgacttcaacaacttgcgttcaaatltgaagtgtgacgc	2611
Db	4888	GGACACTCCAGAAAGACCTAGTGGAGATCGACTTCATTTGGGGTCAGATCTCAGTGAADAA	4947
Oy	2612	gaacgaaggcgtatcgatgcgaacgaaccccttgcgcgtcttacttctgttactacacgctatcgc	2671
Db	4948	GCACCAAGGCTATTCACACCGGACGAGCGCGCGCGCCACGACACTCTCAACCGTGGTGGC	5007
Oy	2672	tacgacgaagatgctccctcttgaaaaacgcgaagaaatctccaagattccaactattctga	2731
Db	5008	CAGGCACAAAGTCCACCTGCTCTCTGTGACGCTGTCTCGCTGATCTCACTACTTTCAA	5067
Oy	2732	agttcacggcgatatacgaaatagaataatcatcgacacaaacaaagaatctcgtgaacgcgt	2791
Db	5068	CACGCCGGAGAGATATACGATTCACATTAACAGACGAAACAAATTAAGTGTCCGTGAGTGA	5127
Oy	2792	cactggaagaattcaaacacgacgcgtgtgacgtcttgcgtgaagaaatgcccaagacgcgcgaagct	2851
Db	5128	AGAGGACAAAGTTTCAGAGAAACTGTCTATATCTCGAGGGCGCGGAGGAGCGGAACACGT	5187
Oy	2852	ggtccagctgtacctcccaacgaacgttgaacaaagacgagttgtgacacaaacgatacgtataac	2911
Db	5188	CGTGCAACTCTTCCCGCAGCATCTGGATATAGATGAATCTTACCAAAAGTATGCTTACCA	5247
Oy	2912	gatagaacttgcgttgaaccccgactgcgcggaattctcttcgaggttgcgaactgcgaactgcg	2971
Db	5248	GATCACTAGCGCATCAACCTCGTCTCCGCGACTCTTTCGAGGTATGACTTGGAAACCGG	5307
Oy	2972	tgcacctgaagttgcatcacaccgcgggggaacgaaaaatttgcgaacgcgcatlbggatalgactac	3031
Db	5308	CCTGGTGAAGTCAACAACACTGCCGGGGGAAATCGACAGACGAGGATGAACCCAC	5367
Oy	3032	acatactactcttctgtaattcatatcatgataactcttcttctgtagtgcgtagagaaga	3091
Db	5368	GCATTCGGATCTTCTTCAACCTCATCATATACTTCTATGGAGAAAGAGATGGGCAACCGGAA	5427
Oy	3092	ccgaagacgaagtgtgaatatcttgccttcattgtgatagtgcgaacgaacgcctcgtgaat	3151
Db	5428	TCAGGACGACGACACAAATGTTGGTGGTGGCTGTTGGACATCAACGACCAACTACCCGGA---	5484
Oy	3152	gccactgcgtatgaactccggttctgatatcttcgaagaagagcaagtgtgcgtgtcgtgt	3211
Db	5485	---ACTGCGTAGGGGTCTCTCATGGATATCTGTGAGGGCTGTGTACAGGGGHTGCCGHT	5541
Oy	3212	actcccgaaatctcaagcaacccgaacagggatgaaacgaacacggaacactcgcgtgtcgcg	3271
Db	5542	AAACCCCAATATTTCTTCCGCCCGGACGCGACGACGCGCCGACCGCAACTCCCGGTGTGC	5601
Oy	3272	ttacggaatcccttgaaactctcaagataccgacgaagacatcgcggtgtgcgcgtactcttcaac	3331
Db	5602	GTACGACATGCTCAAGCTCTACGCCCAACGACAGGACATACACTTCTCTCAACTTTTAC	5661
Oy	3332	catgatctgcattgaaa-----acaaaacttgcggaacttgcgaacgctat	3376
Db	5662	CATGATCACCATAGAAAGACAGGGGGCATTCGACCGAGCTGGAGAGGTGGAGACCGCTAT	5721
Oy	3377	ggaacttgaaggggtatctggggcaactcaagaatatcatatcattgaagccttcogaacacgcta	3436
Db	5722	GGATTTAAGAGGCTATTGGGGCACTTATGAAGTACATGTGCAAGCGATATGACCATGTGAGT	5781

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2002, 13:34:47 : Search time 4372.01 seconds
(without alignments)
13513.302 Million cell updates/sec

Title: US-09-715-909-1

Perfect score: 5498
Sequence: ~~caatacaataaagaaga~~.....aaaaaaaaaaactcgag 5498

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST.*
1: em_estfun:*
2: em_esthum:*
3: em_estlin:*
4: em_estpl:*
5: em_estpl:*
6: em_estba:*
7: em_estro:*
8: em_estov:*
9: em_hic:*
10: 9b_estl:*
11: 9b_est2:*
12: 9b_hic:*
13: 9b_gss:*
14: em_gss_fun:*
15: em_gss_hum:*
16: em_gss_inv:*
17: em_gss_pln:*
18: em_gss_pro:*
19: em_gss_rod:*
20: em_gss_vrt:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	62.8	1.1	1100	13	CNS016KD
C 2	54.2	1.0	925	13	CNS0091P
C 3	53.4	1.0	513	10	AA948367
C 4	53.2	1.0	959	11	CG387114
C 5	52.8	1.0	300	11	N98097
C 6	52.8	1.0	656	10	BE602179
C 7	52.2	0.9	314	10	AU198049
C 8	51.8	0.9	1043	13	CNS0145P
C 9	51.6	0.9	523	10	AV401734
C 10	51.4	0.9	519	11	BI140996
C 11	51.4	0.9	704	13	CNS062FV
C 12	51.2	0.9	945	13	CNS01107

C 13	51	0.9	503	10	AL513809
C 14	50.8	0.9	484	11	BF765609
C 15	50.8	0.9	523	11	BF762445
C 16	50.8	0.9	726	11	BG819608
C 17	50.8	0.9	736	11	BG747190
C 18	50.8	0.9	807	10	AU124569
C 19	50.8	0.9	1101	13	CNS002M0
C 20	50.6	0.9	512	10	AV401445
C 21	50.6	0.9	556	10	BE510534
C 22	50.6	0.9	1068	13	CNS04B86
C 23	50.4	0.9	439	10	AV637890
C 24	50.4	0.9	1125	10	AL547503
C 25	50.2	0.9	200	11	BG438325
C 26	50.2	0.9	245	8	BF048950
C 27	50.2	0.9	782	11	BF631315
C 28	50	0.9	195	11	BG156792
C 29	50	0.9	599	13	AQ160671
C 30	50	0.9	910	13	CNS01T91
C 31	50	0.9	935	13	CNS006XK
C 32	50	0.9	1201	13	CNS00Z07
C 33	49.8	0.9	210	10	AU060678
C 34	49.8	0.9	245	10	AU056151
C 35	49.8	0.9	604	13	AA550384
C 36	49.8	0.9	936	13	CNS003RX
C 37	49.8	0.9	957	13	CNS015W7
C 38	49.8	0.9	1184	13	CNS04PAP
C 39	49.6	0.9	510	10	AV405247
C 40	49.6	0.9	512	10	AV402969
C 41	49.6	0.9	700	10	AL507642
C 42	49.4	0.9	500	10	AU087313
C 43	49.4	0.9	927	13	AZ046250
C 44	49.4	0.9	1101	13	CNS00E7Y
C 45	49.4	0.9	1101	13	CNS0180P

ALIGNMENTS

RESULT 1
CNS016KD/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence SP6 end of BAC
BACN16D22 of DrosBAC library from Drosophila melanogaster (fruit
fly) genomic survey sequence.
ACCESSION
AL106855
VERSION
AL106855.1 GI:5624152
KEYWORDS
GSS.
SOURCE
fruit fly.
ORGANISM
Plasmod Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachyera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 1100)
Genoscope.
COMMENT
Direct Submission
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EGGP) -
http://www.edgp.ebi.ac.uk. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.
FEATURES
source
location/Qualifiers
1..1100
/organism="Drosophila melanogaster"
/plasmid="pBelobAC11"
/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACN16D22"

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 15, 2002, 13:41:47 ; Search time 417.54 Seconds

(without alignments)
11288.916 Million cell updates/sec

Title: US-09-715-909-1

Perfect score: 5498
Sequence: 1ccataataaataaagagga.....aaaaaaaaaaactcgag 5498

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Sequenced: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

N_Geneseq_1101:*

- 1: /SID52/gcgdata/geneseq/geneseqn/NA1980.DAT:*
- 2: /SID52/gcgdata/geneseq/geneseqn/NA1981.DAT:*
- 3: /SID52/gcgdata/geneseq/geneseqn/NA1982.DAT:*
- 4: /SID52/gcgdata/geneseq/geneseqn/NA1983.DAT:*
- 5: /SID52/gcgdata/geneseq/geneseqn/NA1984.DAT:*
- 6: /SID52/gcgdata/geneseq/geneseqn/NA1985.DAT:*
- 7: /SID52/gcgdata/geneseq/geneseqn/NA1986.DAT:*
- 8: /SID52/gcgdata/geneseq/geneseqn/NA1987.DAT:*
- 9: /SID52/gcgdata/geneseq/geneseqn/NA1988.DAT:*
- 10: /SID52/gcgdata/geneseq/geneseqn/NA1989.DAT:*
- 11: /SID52/gcgdata/geneseq/geneseqn/NA1990.DAT:*
- 12: /SID52/gcgdata/geneseq/geneseqn/NA1991.DAT:*
- 13: /SID52/gcgdata/geneseq/geneseqn/NA1992.DAT:*
- 14: /SID52/gcgdata/geneseq/geneseqn/NA1993.DAT:*
- 15: /SID52/gcgdata/geneseq/geneseqn/NA1994.DAT:*
- 16: /SID52/gcgdata/geneseq/geneseqn/NA1995.DAT:*
- 17: /SID52/gcgdata/geneseq/geneseqn/NA1996.DAT:*
- 18: /SID52/gcgdata/geneseq/geneseqn/NA1997.DAT:*
- 19: /SID52/gcgdata/geneseq/geneseqn/NA1998.DAT:*
- 20: /SID52/gcgdata/geneseq/geneseqn/NA1999.DAT:*
- 21: /SID52/gcgdata/geneseq/geneseqn/NA2000.DAT:*
- 22: /SID52/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5498	100.0	5498	AAF30933	Ostrinia nubilalis
2	1960.4	35.7	5578	AAV74277	CDNA encoding Mand
3	1826.6	33.2	5527	AAE30934	Heliothis zea Bt t
4	1737.6	31.6	5527	AAE30975	Pink bollworm Bac1
5	1711.6	31.1	5571	AAI15235	Manduca sexta Bac1
6	1547.4	28.1	1604	AAE30940	Ostrinia nubilalis
7	1465.8	26.7	5592	AAE30935	Spodoptera frugiper
8	55.5	1.0	1910	AAQ34349	Streptomyces livid
9	53.6	1.0	789	AAQ10792	Bovine vascular en
10	53.6	1.0	789	AAQ44255	Partial bovine VEG
11	53.2	1.0	1294	AAZ06824	Streptomyces albid

12	53.2	1.0	1294	AAH74537	Nucleotide sequenc
13	52.4	1.0	244	AAE58238	Oligonucleotide D1
14	52	0.9	936	AAE58252	Oligonucleotide D1
15	52	0.9	936	AAE58254	Oligonucleotide D1
16	52	0.9	936	AAE58257	Oligonucleotide D1
17	52	0.9	936	AAE58259	Oligonucleotide D1
18	52	0.9	936	AAE58262	Oligonucleotide D2
19	52	0.9	938	AAE58255	Oligonucleotide D1
20	51.2	0.9	936	AAE58252	Oligonucleotide D1
21	51.2	0.9	936	AAE58254	Oligonucleotide D1
22	51.2	0.9	936	AAE58257	Oligonucleotide D1
23	51.2	0.9	936	AAE58259	Oligonucleotide D1
24	51.2	0.9	936	AAE58262	Oligonucleotide D2
25	51.2	0.9	938	AAE58255	Oligonucleotide D1
26	50.8	0.9	4104	AAQ68997	Human protocadheri
27	50.8	0.9	4104	AAQ68997	Human protocadheri
28	50.8	0.9	4650	AAQ68999	Product of alterna
29	50.8	0.9	4650	AAQ68999	Product of alterna
30	50.8	0.9	10732	AAI10594	Gene encoding a su
31	50.4	0.9	2048	AAE59540	Human secreted pro
32	50.2	0.9	905	AAH74539	Nucleotide sequenc
33	49.6	0.9	244	AAE58238	Oligonucleotide D1
34	49.6	0.9	1421	AAE58238	Oligonucleotide D1
35	49.4	0.9	1354	AAE58245	Zea mays DNA fragm
36	49.2	0.9	5760	AAE50530	Streptomyces livid
37	49.2	0.9	6357	AAE291051	Sequence encoding
38	49	0.9	2940	AAE36689	F. lutescens piper
39	49	0.9	3106	AAA36696	Humanised F-RSV/HN
40	48.6	0.9	905	AAH74541	Nucleotide sequenc
41	48	0.9	696	AAA50834	Zoanthus fluoresce
42	47.4	0.9	1005	AAZ30006	Optimised nucleoti
43	47.4	0.9	1076	AAZ30007	Optimised nucleoti
44	47.4	0.9	1580	AAE61288	Human secreted pro
45	47	0.9	1098	AAH0527	Leishmania infantu

ALIGNMENTS

RESULT 1
AAF30933 standard; cDNA: 5498 BP.

ID	AAF30933	standard; cDNA: 5498 BP.
XX	XX	
AC	AAF30933:	
XX	XX	
DT	09-JUL-2001 (first entry)	
XX	XX	
DE	Ostrinia nubilalis Bt toxin receptor cDNA.	
XX	XX	
KW	European corn borer; ECB; Bacillus thuringiensis; Bt toxin;	
KW	receptor; crystal protein; CryIA; biological control; insecticide;	
KW	crop protection; ss.	
XX	XX	
OS	Ostrinia nubilalis.	
XX	XX	
FT	Key	Location/Qualifiers
FT	CDS	162..5315
FT		/*tag= a
XX	XX	
XX	XX	
PD	25-MAY-2001.	
XX	XX	
PE	17-NOV-2000; 2000OW-US31674.	
XX	XX	
PR	18-NOV-1999; 99US-0166285.	
PR	21-SEP-2000; 2000US-0234099.	
XX	XX	
PA	(PION-) PIONEER HT-BRED INT INC.	
XX	XX	
PT	Flannagan RD, Mathis JP, Meyer TE.	
XX	XX	
DR	WPI; 2001-329223/34.	

QY	1621	acaaacgtacccacgctgagcgagcgctggtgtatagctgaaactgtatcaactgtgaacgagc	1660
Db	1621	acaaacgtacccacgctgagcgagcgctggtgtatagctgaaactgtatcaactgtgaacgagc	1660
QY	1681	aactcccaactcttgaaagaaagaaagcttaccgcgctgcttttaagaaagccgttcgagccg	1740
Db	1681	aactcccaactcttgaaagaaagaaagcttaccgcgctgcttttaagaaagccgttcgagccg	1740
QY	1741	gcttcoccggtgagccagcgctgctccgcttcgacagaaacatacgaaacgactaagtgact	1800
Db	1741	gcttcoccggtgagccagcgctgctccgcttcgacagaaacatacgaaacgactaagtgact	1800
QY	1801	cattgtatgagcgacacgctgctgttgaactacctgttcatagatgaatcaacgagagatcttcg	1860
Db	1801	cattgtatgagcgacacgctgctgttgaactacctgttcatagatgaatcaacgagagatcttcg	1860
QY	1861	tgaagcatgagacgacttgcgcttcgacttaccacccgaaagaaacactattgtttaaagctgcg	1920
Db	1861	tgaagcatgagacgacttgcgcttcgacttaccacccgaaagaaacactattgtttaaagctgcg	1920
QY	1921	ctgaagacatacttggcgacgagcccaacaacaacagctgacccaacccagcctgggtgaataac	1980
Db	1921	ctgaagacatacttggcgacgagcccaacaacaacagctgacccaacccagcctgggtgaataac	1980
QY	1981	tggagagatgttaaaacaacatccctcccaacccctagcttgcctgccttcgactccaagctcg	2040
Db	1981	tggagagatgttaaaacaacatccctcccaacccctagcttgcctgccttcgactccaagctcg	2040
QY	2041	aggagaaagcttcccgaaagagatagagaaatctcccggaatactctctacgaaacccgaa	2100
Db	2041	aggagaaagcttcccgaaagagatagagaaatctcccggaatactctctacgaaacccgaa	2100
QY	2101	ccagagccctacctgtgtgtctcgagatctgagacttggactccaactggcgcaacaaagagga	2160
Db	2101	ccagagccctacctgtgtgtctcgagatctgagacttggactccaactggcgcaacaaagagga	2160
QY	2161	gagagagacaaacccctactggaataagctcgaggtgtatagttatcgaaacgatataccacacg	2220
Db	2161	gagagagacaaacccctactggaataagctcgaggtgtatagttatcgaaacgatataccacacg	2220
QY	2221	aggagcaacccggggtctccgcacatccgagcgctccgtgtgtgcaaaagagatcccggaacaactca	2280
Db	2221	aggagcaacccggggtctccgcacatccgagcgctccgtgtgtgcaaaagagatcccggaacaactca	2280
QY	2281	ccatcgacttcgaggaattcggagagatgcttaccctcaacgctcccggtgagggaaactcaaca	2340
Db	2281	ccatcgacttcgaggaattcggagagatgcttaccctcaacgctcccggtgagggaaactcaaca	2340
QY	2341	ctgtgcatcggagatgagactatgagatgagagcgagcttcaagatacaataatcgaacttgaagc	2400
Db	2341	ctgtgcatcggagatgagactatgagatgagagcgagcttcaagatacaataatcgaacttgaagc	2400
QY	2401	acaaacgcgcacatcttcgcgaaacgagacacgctgaacgacagatgagcggtgagcagactg	2460
Db	2401	acaaacgcgcacatcttcgcgaaacgagacacgctgaacgacagatgagcggtgagcagactg	2460
QY	2461	cgagccagcgacagcttcatctgcgctccggtgcttcgcgacacgaaatcgaaacgcccgtctaca	2520
Db	2461	cgagccagcgacagcttcatctgcgctccggtgcttcgcgacacgaaatcgaaacgcccgtctaca	2520
QY	2521	accaaagctgcctacacatatacaaaacctaagaaaaaacaccccggaagatattgttgaagattg	2580
Db	2521	accaaagctgcctacacatatacaaaacctaagaaaaaacaccccggaagatattgttgaagattg	2580
QY	2581	acttcaacaactggtcacaattgtgaagtgtgaagtgaagacgagcgagctgacgtatcaacaacaccc	2640
Db	2581	acttcaacaactggtcacaattgtgaagtgtgaagtgtgaagacgagcgagctgacgtatcaacaaccc	2640
QY	2641	ggcgctctctactgttactatacaacccgtcatcgcttagcgacgagtgctccctcgtgaaaaacgca	2700
Db	2641	ggcgctctctactgttactatacaacccgtcatcgcttagcgacgagtgctccctcgtgaaaaacgca	2700

QY	2701	cggnaatgtcctccgcagattccaactactctgaaagtcttcagagcatatcgaataagaatacca	2700
Db	2701	cggaaatgtcctccgcagattcccaactactctgaaagtcttcagagcgatatcgaataagaatacca	2700
QY	2761	tcgacacaaacaaacaaagtgccctgagccgcctcaactcgaagaagtctcaacagaaagtgatcg	2820
Db	2761	tcgaacacaaacaaacaaagtgccctgagccgcctcaactcgaagaagtctcaacacgaagtgatcgg	2820
QY	2821	tcctggagaaatgcgacacagagcgcgacgaagtgatgccagactgtctaccocacgaacgtctga	2880
Db	2821	tcctggagaaatgcgacacagagcgcgacgaagtgatgccagactgtctaccocacgaacgtctga	2880
QY	2881	gagacagagtgtgtaccacacggtatcagataacagatgaactttgcggttgaaaccccgatcgc	2940
Db	2881	gagacagagtgtgtaccacacacggtatcagataacagatgaactttgcggttgaaaccccgatcgc	2940
QY	2941	ggagattctctcgaggtgagacctctgagacactgcgtccgccttgagtgatattacccggggagcg	3000
Db	2941	ggagattctctcgaggtgagacctctgagacactgcgtccgccttgagtgatattacccggggagcg	3000
QY	3001	aaaaatctgacacgagatctgaggagatcgagactcaacatactactctcttgtaaatltcatcgata	3060
Db	3001	aaaaatctgagacacgagatctgaggagatcgagactcaacatactactctcttgtaaatltcatcgata	3060
QY	3061	actctcttctctgtatgtgtacggtctagagaaacacagagacgaagtctgaataattgtcgcttc	3120
Db	3061	actctcttctctgtatgtgtacggtctagagaaacacagagacgaagtctgaataattgtcgcttc	3120
QY	3121	tattggatgtggaagagaacaaagcctccctgagatgcatctgcatctgaacttcacgttttgatg	3180
Db	3121	tattggatgtggaagagaacaaacgctccctgagatgcatctgcatctggaacttcacgttttgatg	3180
QY	3181	ttctccgaagagagcagatgtgcgtgtcgtgtccgtgtactcccaaaaaatctacagacacggagag	3240
Db	3181	ttctccgaagagagcagatgtgcgtgtcgtgtactcccaaaaaatctacagacacggagag	3240
QY	3241	atgaaacacagaaacagagacaactcgcgtgtcgtgtacggaatctctgagcctcaacatcacag	3300
Db	3241	atgaaacacagaaacagagacaactcgcgtgtcgtgtacggaatctctgagcctcaacatcacag	3300
QY	3301	accgagagcatctcgaggtgtgcgcgatactcttcaacatgatactcgaattgaaaaacaaactgggg	3360
Db	3301	accgagagcatctcgaggtgtgcgcgatactcttcaacatgatactcgaattgaaaaacaaactgggg	3360
QY	3361	aaacttgagacgcgtactatgaaactctgaaaggggtatgtgggacactacgaacaaatctcatctggag	3420
Db	3361	aaacttgagacgcgtactatgaaactctgaaaggggtatgtgggacactacgaacaaatctcatctggag	3420
QY	3421	ccttcgcacacagcgtctaccgcgacgacgaaggtctcaacacgagagcttacaacctgtgtatccgc	3480
Db	3421	ccttcgcacacagcgtctaccgcgacgacgaaggtctcaacacgagagcttacaacctgtgtatccgc	3480
QY	3481	ccttcaaacattccacacacacccgtgttggtgtgttcccggaacccgaacccgtgtatccgcttt	3540
Db	3481	ccttcaaacattccacacacacccgtgtgttccgtgttcccggaacccgaacccgtgtatccgcttt	3540
QY	3541	ctaaagagcgcgcaacacgaagagcgcggtctctctgcgacagcgtctccaaacagatctctctggagc	3600
Db	3541	ctaaagagcgcgcgcaacacgaagagcgcggtctctctgcgacagcgtctccaaacagatctctctggagc	3600
QY	3601	cgaatctacgcacacacgacgagagcgagcctccacacgcgagcgatcgaagtcttccagctccag	3660
Db	3601	cgaatctacgcacacacgacgagagcgagcctccacacgcgagcgatcgaagtcttccagctccag	3660
QY	3661	gaaatgagggggcggtttcagttacttggatataatactgaagtggtggagcgagggaaaaatagcg	3720
Db	3661	gaaatgagggggcggtttcagttacttggatataatactgaagtggtggagcgagggaaaaatagcg	3720
QY	3721	ggcagcgtataattacgacacagcgttttcccgagagcaaatcatgacaattcagatcagctacc	3780
Db	3721	ggcagcgtataattacgacacagcgttttcccgagagcaaatcatgacaattcagatcagctacc	3780
QY	3781	gggcacacagagcggcgacaggaaccccgagccgcttttgagacacgaatcagcttttccgtatgg	3840

```

Db 3761 gggccacgagcggcgccgagccggccgcttggaccgacgttcggttggctg 3840
Oy 3841 tcttgcatacccaacagagggacccagtgcttaagcgaataatgcaactgtgcctct 3900
Db 3841 tcttgcatacccaacagagggacccagtgcttaagcgaataatgcaactgtgcctct 3900
Oy 3901 tcgaggtgtaagaagggctccgttgagagtttgagctgcccgaacgaagaagccttaaa 3960
Db 3901 tcgaggtgtaagaagggctccgttgagagtttgagctgcccgaacgaagaagccttaaa 3960
Oy 3961 accaactctgcaagaatgactgccaagatactactacaggtttatgagcgcaaacg 4020
Db 3961 accaactctgcaagaatgactgccaagatactactacaggtttatgagcgcaaacg 4020
Oy 4021 agggccttctgacttgaccaggttaacgaagatcatctcccttcgagaggttggac 4080
Db 4021 agggccttctgacttgaccaggttaacgaagatcatctcccttcgagaggttggac 4080
Oy 4081 gcgaggtgagcagcttctacacgctgacacatcgccgagcaactcgccgacgacatg 4140
Db 4081 gcgaggtgagcagcttctacacgctgacacatcgccgagcaactcgccgacgacatg 4140
Oy 4141 ggaatctctgcagagcttccatctctgtgtcagcgtcaatgtaagaagaagcgaacccg 4200
Db 4141 ggaatctctgcagagcttccatctctgtgtcagcgtcaatgtaagaagaagcgaacccg 4200
Oy 4201 gcccaattctgagagagaccttaccagacgagcttgcagcttgcagagcttgcagc 4260
Db 4201 gcccaattctgagagagaccttaccagacgagcttgcagcttgcagagcttgcagc 4260
Oy 4261 gggaaattgacttctgctcagggcgagccacacagaagacgacacca tcaagtaaccatag 4320
Db 4261 gggaaattgacttctgctcagggcgagccacacagaagacgacacca tcaagtaaccatag 4320
Oy 4321 accgtgagagcagctgagctggaacgacgactagaagccgtgacgactcgccctcgcc 4380
Db 4321 accgtgagagcagctgagctggaacgacgactagaagccgtgacgactcgccctcgcc 4380
Oy 4381 tgcagcagcagcggcgctgcttctgctcaataatgacgcccacgcttccatgcagcga 4440
Db 4381 tgcagcagcagcggcgctgcttctgctcaataatgacgcccacgcttccatgcagcga 4440
Oy 4441 tgttcagctcagcgtlcatcgctacggtacagcttctgcaatgcagacagccgtgtga 4500
Db 4441 tgttcagctcagcgtlcatcgctacggtacagcttctgcaatgcagacagccgtgtga 4500
Oy 4501 aagttactctatctcactcgcgaacacccgctgacacttcaatttcgataaccaacttga 4560
Db 4501 aagttactctatctcactcgcgaacacccgctgacacttcaatttcgataaccaacttga 4560
Oy 4561 ccgtgagcagagacgaatattcagcggcagcttcagcagccggttcaacatgagct 4620
Db 4561 ccgtgagcagagacgaatattcagcggcagcttcagcagccggttcaacatgagct 4620
Oy 4621 gcaacatcgaacaggtggtgcgttcagcagcagcagcggtgagcgaagaagcgaacca 4680
Db 4621 gcaacatcgaacaggtggtgcgttcagcagcagcagcggtgagcgaagaagcgaacca 4680
Oy 4681 ccgaggtgagcggcagcttcaatccggggaacagctgcccgttgcaagcacaagaggtcga 4740
Db 4681 ccgaggtgagcggcagcttcaatccggggaacagctgcccgttgcaagcacaagaggtcga 4740
Oy 4741 ccgtcgcagcagcagcgtgtgctgctgcacacatccagctgtgtgacacacacagcc 4800
Db 4741 ccgtcgcagcagcagcgtgtgctgctgcacacatccagctgtgtgacacacacagcc 4800
Oy 4801 tgggtctgcaagacctgtgtgacggttgacactcgcagcgttaggcgagagatcagcaga 4860
Db 4801 tgggtctgcaagacctgtgtgacggttgacactcgcagcgttaggcgagagatcagcaga 4860
Oy 4861 tcgctgcttaagcattacgagcgtctccgctgtgtgaggttctctgctctgactgcg 4920

```

```

Db 4861 tcgctgcttaagcattacgagcgtctccgctgtgtgaggttctctgctctgactgc 4920
Oy 4921 tcttcgcatgttctctgttagacaagaagcactgtaaccggaagcagctgcaagcacttccatga 4980
Db 4921 tcttcgcatgttctctgttagacaagaagcactgtaaccggaagcagctgcaagcacttccatga 4980
Oy 4981 cgaagtlacgctcgtgtgactccgggctgaaacgcgcgcggcgctgagcgcgggacccaaca 5040
Db 4981 cgaagtlacgctcgtgtgactccgggctgaaacgcgcgcggcgctgagcgcgggacccaaca 5040
Oy 5041 agcaagcgtctgagagggctccaaacccatgttggaagagccatccgcgcgcgcgaatttcg 5100
Db 5041 agcaagcgtctgagagggctccaaacccatgttggaagagccatccgcgcgcgcgaatttcg 5100
Oy 5101 agccatcagtgagcggagtgcgactccgactccgactatcgcatcgaagacatgcccgaat 5160
Db 5101 agccatcagtgagcggagtgcgactccgactccgactatcgcatcgaagacatgcccgaat 5160
Oy 5161 tcgagcagcactactccgcgcgcgcggaacagactcgaagcagcgcatcgtctgcaca 5220
Db 5161 tcgagcagcactactccgcgcgcgcggaacagactcgaagcagcgcatcgtctgcaca 5220
Oy 5221 tggcggaagccacagcagcagcccgtagacacgcatggaacacacttgggttcaagt 5280
Db 5221 tggcggaagccacagcagcagcccgtagacacgcatggaacacacttgggttcaagt 5280
Oy 5281 ccaccgcgtactctgcacagccgaccccaagtaactgcaggggtataacctgttccagg 5340
Db 5281 ccaccgcgtactctgcacagccgaccccaagtaactgcaggggtataacctgttccagg 5340
Oy 5341 tgcctagccgcgcggaagtgcgcacagcgttattcatcgcgggaacattagatgaagat 5400
Db 5341 tgcctagccgcgcggaagtgcgcacagcgttattcatcgcgggaacattagatgaagat 5400
Oy 5401 acctatgacatattgttaattgttaacatcattatattataacatattatattat 5460
Db 5401 acctatgacatattgttaattgttaacatcattatattataacatattatattat 5460
Oy 5461 attgctaaaaaataaaaaaataaaactcgag 5498
Db 5461 attgctaaaaaataaaaaaataaaactcgag 5498

RESULT 2
AAV74277
ID AAV74277 standard; cDNA: 5578 BP.
XX
AC AAV74277;
XX
DT 05-JUL-1999 (first entry)
XX
DE cDNA encoding Manduca sexta BT toxin receptor BT-R1.
XX
KW BT-R1; receptor; tobacco hornworm; Bacillus thuringiensis;
KW BT-toxin; cadherin; insecticide; pesticide; ds.
XX
OS Manduca sexta.
XX
FH Key
FH CDS 197..4783
FT /*tag= a
FT /gene= "BT-R1"
FT sig_peptide 197..259
FT /*tag= b
FT mat_peptide 260..4780
FT /*tag= c
FT polyA_signal 5564..5569
FT /*tag= d
FT misc_difference 4989..4991
FT /*tag= e
FT /*note= "these bases represent nucleotides missing
FT for the sequence given in the
FT specification. They are included to

```


QY	3722	gagcttatataagccagctttccagagcaacatcagacaattccagatcacatccg	3701
Db	3769	ctctgtgaggttacctgcgaacgcgttccagaggagatcagagggtttccggtataacgattccg	3828
QY	3782	ggccacggagacgcgccgacgcgagccgcgcctcttgacacgcagcttccglttcgtgtg	3841
Db	3829	cgtataagcaccagggaacgcgaccacagaccgcgtctccacgagcatgacgtttcaggtgt	3888
QY	3842	cttctgatacccaacaaggcgacccagtgcttcagcgaaaatgcagctactgttcgctctt	3901
Db	3889	ttttgtgccccacgcgaaggagaaactcagattcgcgtctctacagacaatgcttgcgtttat	3948
QY	3902	cgaagggtgaagaaagcctccgcttcgtagagttttgagctgcgcgcgacaggaagactataaa	3961
Db	3949	agaaaaagaggttcgcatgtggaagaagttccaccaacttctcttagacaagatacaaa	4008
QY	3962	ccacactctgcgaagaatgacgtgcgaagaatatctatacagtttatgtgacgcgaacaaga	4021
Db	4009	ccaatctcgtgaagaacgactgtccacagcatcttatactgtattatgatggaacaagga	4068
QY	4022	gggtcttttctgtactgtgacacagtcgaagcaacttcatctccctgcgcgaggttgagccg	4081
Db	4069	aggtcaattctgcgcctgtgatactcgttctgcgaacaagttgttctctggaagaagaagctgataag	4128
QY	4082	cgaagtgccacgcgtcttaccacgcgtccacatcgcgcgcgaggaagcaactccgcgaagccagtg	4141
Db	4129	ggaaacaaggtgcctccccaacactctgcgaagtgtggcggtatgtaactgcgcgaatg---	4185
QY	4142	gataccctctgcgaactccatccctcgttgcatacgttcaatgttaagagaagcgaaaccgcg	4201
Db	4186	catccacacttccgtctccatctccatcttactcactgttaccggtgaaggaggagaaaccctcg	4245
QY	4202	cccaattctcgagaaggaaccttaccacagcgggcatcttgaaagtttgacagcatgtgcgcg	4261
Db	4246	tccagttgtttgtgagggaaattgtaccacgcgcgaggaatccaaaccagcgcgactccatccgcgaag	4305
QY	4262	ggaatgtcttactgtctcaggggcgcgcacacagagaagacacacalcacatgcataccatat	4321
Db	4306	agagctgtctcagattacatgacgcacccagctcgaaggtctgcgcatacttactatgataga	4365
QY	4322	ccgtgcgcgagcatgcagctgtgacagcagcgccttagaagccgtgtgcgactgcgcctcgcgt	4381
Db	4366	ctacgataacatgtgtatgtgaccccacgcgcctgtgaaggagatgaaacagctgcgtttcgtact	4425
QY	4382	gcatgcgacccacgcgcgtgttcttcgcctcaatatgcagccacgcgcttccatgcagacgat	4441
Db	4426	gaaagcttcaaacgcggaggtgcgttcgcttaatatccagccacgcgcgcacagaatgcattgaa	4485
QY	4442	gttcgaagtcgaogtcaatcgcatacgaataccttcgcaatcgcacacagcccgltgaa	4501
Db	4486	gttcaaatctcgaagtgtgacagactctacacacgcgcgcgtctcagagccgcacgcgaagctaac	4545
QY	4502	agttactactatctcattcgcgaaaaacgcgtgtgactttcatttctgtataccaacttttgac	4561
Db	4546	cgtgtacgtgtgatactcctgcgaagaacgcgcgtcttactcgtgtgtcgttcaaacacgcgtgcaca	4605
QY	4562	cgttgagcagaacagaatattctaaatgcgcgcacgttccagcgcacgcggtttcaacatgacgtg	4621
Db	4606	ggtcgaagacaacagagactttatctgcgaacaccttcaagcgtctggtgttcaacatgacctg	4665
QY	4622	caaatctgcgaaggtgtgtgcgttccagga---cagcagcgcgcgtgtgcgcgaagccgacac	4678
Db	4666	caaatctgcgaaggtgtgtgcgcgttccagtaacgaaccgcgttcaacgcgcgtgcgcgtgcgcgcacag	4725
QY	4679	caccgcgtgtgcgcgcgcacattctaccgggagacaacgtctccgcgtgcgcgcgcagacaggttgcga	4738
Db	4726	cacgcagatgt-gcgcgcacattctacacgggaacaacgttaccgcgttactcgtgtgtgatataga	4784
QY	4739	ggcgcgttccgcgcgcgcagacggtgtgtgtcgtgcgcacatccagctgtatgctgtgacccaag	4798
Db	4785	acagatccgtgtgtgtacctatgcttctctgtgcctgcgataaacaacagcgtg-gcgcgcgat	4843

OY	4799	ccggtctcgtaagaccttggtagacgggtgaacaactccgaacgcataaggagatcaatgca	4858
Db	4844	cgtgtcttgtaagactttgtttccaactcagaaccggagattccggtcatctgaac-	4902
OY	4859	gatgcgcgtactaacgaattaacgcgcgtcttcgcgtgtgcatgcttcctctgcctaag	4918
Db	4903	-ctgcaacgggtgaagtcttgctccactgtctcgtgtgcgcggtttccatggtcgttgtct	4961
OY	4919	gctcttcgcgattgttctctgtaggaacaagacatctgaacgcgcagctgtcaagcata	4978
Db	4962	actgcttaactcttatcatcactagactaginnmgctaaacgcagcgtgtggaagcccctgcat	5021
OY	4979	gacgaagtacgcgcctcgtagaacctccgggcctgaaacgcgcgcgggtc---ggcgcggggac	5035
Db	5022	gacgaagtacgcgcctcacctgtgactcttgatgtgaacgcgcgcgggcatacgcgccccggac	5081
OY	5036	caacaagaacgcgcctcgtagagctccaaacccatgttgaaacgggggcacatccgcgcgcga	5095
Db	5082	caacaagaacacctgtggaaggtcccaacctcttccaattgaagaacaataagaacgcga	5141
OY	5096	cttcgaacacatcaatgtgacgcgaatgtggaactccgaactcgaatcgtgcatactgaagatatcc	5155
Db	5142	ttagatgtccattagcgaaggtctccaacgactctgctgtacatcgcatacgaagaattctgc	5201
OY	5156	gcaattccgcgaacgacatacttccccgcgcgcgaacaaga	5194
Db	5202	gcactcttggaacgcttctcatgtatccttgaagtgaacga	5240

RESULT	3
AAF30934	
ID	AAF30934 standard; cDNA; 5527 BP.
XX	
AC	AAF30934;
XX	
DT	09-JUL-2001 (first entry)
XX	
DE	Heliothis zea Bt toxin receptor cDNA.
XX	
KW	Corn earworm; CFW; Bacillus thuringiensis; Bt toxin;
KW	receptor; crystal protein; CryIa; biological control; insecticide;
KW	crop protection; ss.
XX	
OS	Heliothis zea.
XX	
EH	Key
FT	Location/Qualifiers
CDS	162..515
FT	/*lag= a
XX	
PN	W0200136639-A2.
XX	
PD	25-MAY-2001.
XX	
EF	17-NOV-2000; 2000MO-US31674.
XX	
PR	18-NOV-1999; 99US-0166285.
PR	21-SEP-2000; 2000US-0234099.
XX	
PA	(PION-) PIONEER HI-BRED INT INC.
XX	
PI	Flanagan RD, Mathis JP, Meyer TE;
XX	
DR	WPI; 2001-329223/34.
DR	P-PSDB: AAB20499.
XX	
PT	New nucleic acid encoding Bt toxin receptor, useful for screening and
PT	identifying new Bt toxin receptor ligands useful as new insecticidal
PT	toxins -
XX	
PS	.Claim 1(a); Page 61-66; 85pp; English.
XX	
CC	The present sequence of is that of cDNA encoding the Bacillus
CC	thuringiensis (Bt) toxin receptor (see AAB20499) of the corn

[illegible][illegible]

[illegible]

XX	AAAF30975	4	RESULT
XX	ID	AAF30975	standard; cDNA; 5527 BP.
XX	AC	AAF30975;	
XX	DT	23-JUL-2001	(first entry)
XX	DE	Pink bollworm <i>Bacillus thuringiensis</i> toxin receptor BT-R2 cDNA.	
XX	KW	Pink bollworm; <i>Bacillus thuringiensis</i> ; crystal toxin; CryIac;	
XX	KW	receptor; BT-R2; insecticide; pesticide; biological control;	
XX	XX	cotton; ss.	
OS		<i>Pectinophora gossypiella</i> .	
XX	Key	Location/Qualifiers	
FT	CDS	154..5343	
FT		/*tag- a	
FT	sig_peptide	154..213	
FT		/*tag- b	
FT		/note= "The signal peptide is alternatively	
FT		encoded by nucleotides 154..222"	
FT	mat_peptide	214..5340	
FT		/*tag- c	
FT		/note= "the mature protein is alternatively	
FT		encoded by nucleotides 223..5340"	
XX		WO200134807-A2.	
XX	PD	17-MAY-2001.	
XX		25-OCT-2000; 2000MO-US41521.	
XX	PF	26-OCT-1999; 99US-0161564.	
XX	PR	24-OCT-2000; 2000US-0161564.	
XX	PA	(BULL/) BULLA L A.	
XX	PA	(CAND/) CANDAS M.	
PI		Bulla LA, Candam M;	
DR		WPI: 2001-335932/35.	
XX		P-PSDE; AAB82331.	
XX		Novel <i>Bacillus thuringiensis</i> toxin receptor from pink bollworm, for	
PT		identification and design of toxin receptors, for use in combating	
PT		emergence of toxin resistance and producing transgenic organisms -	
XX		Claim 3; Fig 1A-B; 56pp; English.	
XX		The present sequence is that of cDNA encoding a novel 194 kDa high	
CC		affinity receptor, termed BT-R2 (see AAB82331), of <i>Bacillus</i>	
CC		<i>thuringiensis</i> (Bt) crystal toxin CryIAC. The cDNA was obtained	
CC		from pink bollworm (<i>Pectinophora gossypiella</i>), the most destructive	
CC		pest of cotton worldwide. Degenerate primers based on conserved	
CC		regions of <i>Manduca sexta</i> and <i>Bombyx mori</i> BT-R1/5 were used to clone	
CC		partial fragments of BT-R2 from insect midgut RNA. Based on the	
CC		sequences of the partial clones, sense and antisense primers were	
CC		designed and used to clone the 3' and 5' ends of the BT-R2 clone by	
CC		RACE. The full-length cDNA was then obtained using gene-specific	
CC		primers from the 5' and 3' untranslated regions. The minimum toxin	
CC		binding fragment of BT-R2 has been identified. Isolation of BT-R2	
CC		cDNA permits the analysis of receptors in pink bollworm and other	
CC		insects that affect crop growth and development, as well as the	
CC		development of assays for the cytotoxicity and binding affinity of	
CC		natural and/or introduced homologous receptors and hence the	
CC		specific destruction of organisms, tissues and/or cells of a target	
CC		host, including insects resistant to Bt toxins. Transgenic	
CC		organisms expressing toxin receptors can also be generated. The	
CC		BT-R2 cDNA sequences will also facilitate understanding of the	
CC		molecular biology of the toxin receptors in the pink bollworm and	

CC the engineering of more effective toxins in terms of longer
CC persistence in the field and higher toxicity.

Sequence 5527 BP, 1550 A; 1340 C; 1371 G; 1266 T; 0 other;

Query Match 31.6%; Score 1737.6; DB 22; Length 5527;

Best Local Similarity 61.5%; Pred No. 0; Mismatches 1804; Indels 42; Gaps 9;

Matches 2952; Conservative 0; Mismatches 1804; Indels 42; Gaps 9;

QY 388 tcaatatacagaggaagaatcgaagagacgtacatctgcaagcttaactcaag 447
DB 410 tcatcacatgaagagagagcgagagatgtgtctatcgctaacttaactatgag 469
QY 448 ggtccaaacgcgcgtgtgtctgtatgtcagcgccagagccagagcggtgagccgt 507
DB 470 gttctggaaccccaaaatgtccagccgatgttataagatcttcaacgtcctaagtc 529
QY 508 agtttcgagagatgaagagcgccaaatgagagcgtgttatttcgcaagaagagact 567
DB 530 cagagatccggaatga---aaacggggcggtgttacccttataaacaataagcaagatt 586
QY 568 acgagacagcaacacatgcaagactatgtgttcacatccaagtgagggatgatacag 627
DB 587 atgaaacacacacatgctgtgtatatacttcagcgtccgagtgccagagcagactcgt 646
QY 628 ccgtactgtgtgctgtgagatagtcacatcgaagacaaaccgccactcctgcaagt 687
DB 647 cggcgcagatgtgtctgttcacgaacaaatgcaagatacgaacctatcgtcgaagttgc 706
QY 688 tcaagcgcctgttaattcccaagacatgcaagctagaactgacagcagctgtgataag 747
DB 707 tagaagcgttgcaagatgcccgaattggggagacccctgcactaacaacagcgtttacc 766
QY 748 tgtcagaccgcaagcgttgaatcgaaccccgcttcagacgttcgtgtcgaacagca 807
DB 767 tgcacagcaagaatgtggaggttaagtatcagccatgacatccctcactacatagacc 826
QY 808 gggcgtcagatgaagactcttctacatgtgttgagatagaccccgagcgtgttca 867
DB 827 g---tgaagacgtlaagaaatctctatgtgagcagctacatcactgtgtgtgttca 883
QY 868 atatgaagatgactgtggagatcaatctcccttgactcgaagacacatcagctacata 927
DB 884 acatgcaaatctactatcgttatcctactacagcgttaacttcgaagaacacccgctgaca 943
QY 928 tattaagctcaagactctgactcgtactcgaacaaacacacagcgttcaactatgtgtgc 987
DB 944 tcttcaaatcactgtgtgtgactctcgtccacacacacacacaggtgtgagtggtgc 1003
QY 988 aagtgagagagctagagctcgtccctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1047
DB 1004 aagtcacgaatgtggaacacgcgcgcgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1063
QY 1048 ttgacgagagacatcaatcagactctccctcgcgcgcgtgtgtgtgtgtgtgtgtgtgt 1107
DB 1064 ttgacgagatgagcagacgaatctccaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1123
QY 1108 atagagcacaactatccctacatcagagatgagcgtgtgacgaactcttccctgga-- 1165
DB 1124 ggaagacatataactatccctcgcgcgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1183
QY 1166 -----ggtgattgaagagagctattctgtcagctgtgactgtgagtgacggcgcaagc 1218
DB 1184 caattccggcgccgcaatgagagacacatcttcaagcactgtcagatgtatgtgtgtgtgt 1243
QY 1219 ttgaaagagagcttttcaactcaccatcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1278
DB 1244 tccggcagatgtcttcaagcgtgtccctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1303
QY 1279 caacagagccacatcttcatcatcgttcaacagcgtcaatgtatcagcagccagccgc 1338
DB 1304 ccaaccgcaacccgtcgtgtatcatagatgacatcaacaagaacacccacccgc 1363

QY 1339 tgcataaagaatacagatattgtatcatcgtgaggaactccaatgtactcaactcaatg 1398
DB 1364 tgaagatgtatgatacaaatctccatattgtgaagaactccactgtcgtgtattgtgtc 1423
QY 1399 aagaatttgattccatgtatcgtgagatgtgtgtgaagaactccactcaacagatgagact 1458
DB 1424 aacttttggttctcatgtatgaagattg---atctagcacaatccctgtgtgtgaatc 1480
QY 1459 aggaagtggtcccgcaagggcggtccgcacatctacatcgcgcggggagcggtacc 1518
DB 1481 aagcgagagacccctccagcggtgagcaagcggttataattgtccacacagagcttcc 1540
QY 1519 agagcagaaccttcatcatgt 1578
DB 1541 agaacacgacatctgcataatgagatcgaagatcccggaatgtgtgtgtgtgtgtgt 1600
QY 1579 ttttcagagacatcatcattgaagtcgaagcagtgagacatgcaacagctagccagctgt 1638
DB 1601 ctttccaaacatcaagctcgaaggttaataagcaacgagccgtgtacataacatttactg 1660
QY 1639 gcgaagcgctgt 1698
DB 1661 ggtcgcggaagtcgaagcgtgaactgtatattgtgaacgagcagggagcgcgactttgag 1720
QY 1699 agagcagctactccgcgt 1758
DB 1721 aagacacagctgt 1780
QY 1759 tgcgcctcctcagagacatcagcagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1818
DB 1781 tgaaggttcacagaccgggagacatagagacagcgtgtgtgtgtgtgtgtgtgtgtgtgt 1840
QY 1819 ttgactactgttcatagatgaatcaacagggagagactcgtgtgtgtgtgtgtgtgtgt 1878
DB 1841 atacatttttgagatcagcaggaagaactgvcgacatagctagctattgtgtgtgtgt 1900
QY 1879 tgcactacacccagagaaacactattgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1938
DB 1901 tgcatttcaacagagaaatgaatttaacatagaagttcgtcgttcaaggaacacagctgt 1960
QY 1939 a---cggcccaacacacacagtgacacacagcgtgtgtgtgtgtgtgtgtgtgtgtgt 1995
DB 1961 agccagagttccagagatcagcagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2020
QY 1996 acaactcctccac 2055
DB 2021 ac 2080
QY 2056 aaggaatgagatcccggaatcaactgtctacccgacccggaacacacacacacacacac 2115
DB 2081 agggcttgaatcaacacgggagataacccgacacacacacacacacacacacacacac 2140
QY 2116 ggttcgagatcagactgtgagctcagctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2175
DB 2141 agttggaatgattgtggacacatccttgcacacacacacacacacacacacacacac 2200
QY 2176 ctgaatcgtcgt 2235
DB 2201 tagagttccagagatgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2260
QY 2236 ccgcacatcgcgcgt 2295
DB 2261 aggcgt 2320
QY 2296 aatcgaatgcttcaactcagcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2355
DB 2321 agttggaatttctcactcaacaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2380
QY 2356 actacgagtgaagcgaacttcaacgacatcaacatcaatcgaatgaacgaacacgcgcgact 2415
DB 2381 attatgaaatctactcctcaggaatcaataatagatagatgaacgaacacacgcgcact 2440


```

Db 4595 ttgagagtaacagagacttatatgcagaaacglttcagcgtttggttcaaatgacctgca 4654
Oy 4624 acatcaccaggtgtgtccgttcagcgacagcagcgccgttgccgagcaacaccacccg 4683
Pt 4655 atatagataggtgtgtccggtgacccaagcagcgccggtgtatctcagagacgtgtgcg 4714
Oy 4684 aggtgcgcgcacttcactcaccggtgacaaacgttccgttgcagacaaagagtgacgccc 4743
Db 4715 aagtcacgtctcacttcatacaagataaataccctctgtgagccgcaglatgtgaagagc 4774
Oy 4744 tccgcagcgacacagtgctgtcgtgcgcacacatccagctgtatgtctgagaccaaacgctcg 4803
Db 4775 ttgcgcagtgacactcagctcgtcgtccgtcccaaggttgttgaacccaacggtctgttg 4834
Oy 4804 tgcgtgaagacacgtgtgagcgggtgacacatccgcagctgtagcgagagtgtaatgacatcg 4863
Db 4835 tctcgaagacactgtgtgacgggggttcagccctgtatctgcgcactcgcgcggtgcgacatca 4894
Oy 4864 ccgtctacgcactagcgcgcgtctccgtctgtctgaagcttccctcgtcgtactgtctc 4923
Db 4895 ccatctatgtcagccgggttgtagcagcatccctcgtcctcgtctcgtatcttctgca 4954
Oy 4924 tgcgatgtctctgttagaagaagacactgaaccgagcagctgcagacactctccatgacga 4983
Db 4955 tcaacttcactgttgagagaccgactctgaaccgcttggaaagacactgtcgtatgacga 5014
Oy 4984 agtacgcgtcgtgtgacatccgcgctgaaccgctgacggcggtctgagcgccgacccaaca 5040
Db 5015 aatacgcgtcgtgtgacatccgcgctgaaccgctgacggcggtctgagcgccgacccaaca 5074
Oy 5041 agcacgcgtcgtgaggtctccaaaccccatgttggaacgagccatccgcgcgcgcgactcg 5100
Db 5075 aacacgcctcgtgaaggtctcaaaccccatctggaacgagcagatcaagcgcccgactcg 5134
Oy 5101 accgcatacgtgacgagtgagcagctccgcagctgacatcgcatcgagacatgcgcga 5158
Db 5135 atgcacatcagtgacacatctgacactgtgacttaatcgtgacatcgagatgactgca 5192

```

RESULT 5

AAAT15235
ID AAAT15235 standard; cDNA; 5571 BP.

AC AAAT15235;

DE 10-NOV-1996 (first entry)

XX Manduca sexta *Bacillus thuringiensis* BT-R1 toxin receptor cDNA.

KN Tobacco hornworm; crystal protein; pesticide; insecticide; toxin;
KM cryIA(b); cadherin motif; pest resistance; insect midgut; ds.

XX Manduca sexta.

OS Key Location/Qualifiers

FN CDS 197..4783

FT polyA_signal 5561

FT /*tag= a

FT /*tag= b

XX WO9612964-A1.

XX PD 02-MAY-1996.

XX PF 10-OCT-1995; 95WO-US13256.

XX PR 19-OCT-1994; 94US-0326117.

XX PA (UNIV -) UNIV WYOMING.

XX PI Buila LA, Ji TH;

XX DR WPI; 1996-230725/23.

```

DR P-PSDB; AAR95333.
XX
XX New nucleic acid encoding receptor for Bacillus thuringiensis toxin
PT - useful for treating malignancies, metastases and infections, and
PT for screening of pesticides
XX
XX Claim 1; Fig.1a-1i; 33pp; English.
XX
XX This sequence may be expressed recombinantly in COS-7 or embryonic
CC 293 cells, after cloning in plasmid pCDNA3. The encoded protein, BT-
CC RI specifically recognises the cryIA(b) toxin of B. thuringiensis
CC berliner, which selectively kills M. sexta larvae. The receptor may
CC be used in assays for candidate pesticides (especially insecticidal
CC toxins). Transformed cells, particularly where BT-R1 is expressed
CC at the surface, is used to assess binding affinity/cytotoxicity of
CC candidate pesticides.
XX
XX Sequence 5571 BP; 1483 A; 1542 C; 1337 G; 1209 T; 0 other;
SQ

```

Query Match 31.1%; Score 1711.6; DB 17; Length 5571;
Best Local Similarity 63.1%; Pred. No. 0; Mismatches 1579; Indels 129; Gaps 13;
Matches 2918; Conservative 0;

```

Oy 555 caaagacaagactacagacagacaacacatgcagagctatgtgttcaatccaagtggag 614
Db 674 ccaacaacacatatacagatgtgcgcccaacacacacacatatagttcaatgtgcgtggac 733
Oy 615 ggtgaacacagccgtactgtgtgcgttgagatagtcaacatcgcagacaatccgcc 674
Db 734 ggcgcgtcgtgtgtgacgctgtctcgtatcgtcaacatagatgacaacgccc 793
Oy 675 atccgcgaagtgtgaagcgtcgttaattccagacaatggcgaggttagactgacgac 734
Db 794 atccatcaaaacttcgaagccttgcgggttccctgaacttggcgagccagggttgacaga 853
Oy 735 tgcgttacaaagtgtcaacccgcagcgttgaatcagacacccgcttcaatgcgttcgt 794
Db 854 tgcacatacacaagtatcgacgcgcagcagcagcagcagcagcagcagcagcagcagcagc 913
Oy 795 gtcgaacgacgagcgtgtgcagatgaagaacatctctacatgtgttggaataacgaccc 854
Db 914 atcgacagcgttcg---tgcgacgagagagacctctacatcgaacgagcagatattccc 970
Oy 855 agcgactgttcaatatgaagatgactgtggtgacatctgccttgaacttgcagagaa 914
Db 971 aaccaatgagatgtgttaataatagccatagcgatlaaccctgcctcaacttgcac 1030
Oy 915 actcagctcatatattttagcgtcaacagcttctgactcgtcagcagacaacacacgctc 974
Db 1031 agtcgcgtgcataatatcagcgttgacagccctgcgactcgtctccgaaacccacacggtg 1090
Oy 975 accatgatgtgcaagtgtgagacgtlagatgtcggccctcgtcgtgttgagatcttc 1034
Db 1091 actatgatgtgtgcaagtgtgacagcgttcgcgcgcgctgtcgtgtgagatcttc 1150
Oy 1035 tcaatgacgacatttgaagagaagaactaagagcttccctccgcgcgtgagaaaggg 1094
Db 1151 gctgtccacaagtttgaagagaagaacttcaacaaacttcaacagltgagggcgatcgaga 1210
Oy 1095 gacacgggaatcaatagagccatcaactacacatcagcaggaaggaagcagcagactc 1154
Db 1211 gacacttgatcaatatactcctatcaactacagcgtgacacaaatgaggaagaacacatc 1270
Oy 1155 ttccctcgtgaggtgtg-----tgaagacggagcattctcgtcgtgactgagatc 1205
Db 1271 ttacgacttgagccctcgtcgtgtgaaagaagcggtgttatctcctcgtgtcgcaatt 1330
Oy 1206 gacgcgacaagcttgaagaagagccttcaaccctcaccatcgtgtgttacaactact 1265
Db 1331 gaccgcgacacacttgacaagagaggtgttccacttacgactcgtgttacaataatgat 1390
Oy 1266 gacgttagcttggacaagagagccacatcttcatcatcgttcaacgagcgtcaatgatcag 1325

```

1391 gagagagcccttcacatcaacaacgctggtcatcatctgtgacagacatcaacgaccaa 1450
1326 cgagccggagccgctgtcataagataacagatattgatattctgtgaggaactccaattact 1385
1451 acacccgtgaacctatacaagaagatatacgactgtgcaactatgtagagagccccctgacc 1510
1386 cttaactcaataaagaatttgatccatcgatcgaagatttgggtgtaaaacgtctcaaac 1445
1511 ctcaactccataaagaattcggaattccatgataaagattttagggtccaacgctcagttac 1570
1446 acagttgaaacttgagagacgtgttcccgacagggcggtccgcattctacatcgccg 1505
1571 acggtgtctagagagcgtggaaccccccgcgtctgtagagcattctacatagccgct 1630
1506 ggagagcgctacacagagacgactctatcatgtgacccaataaacacacatcttgat 1565
1631 gaagtcggtacacagcgaacttcatactgcccacccctcaactcacaatgctggat 1690
1566 taagaagatgtcatctttctcaagaacatcatcaataaagtcacaagcagtggaataaac 1625
1691 taagaagtgccagaggttcagagatatacgatctcggtgtgtagcgacacgacacaaacgac 1750
1626 gctagcaacgttggcgagagcgtgtgtgtacgttgaaacctgtaacatggaacgagacact 1685
1751 acgagagcaacgtggcgctgcgtgtgttcaacatccatccatcaattgaaacgattgagcag 1810
1686 cccactctcgagagagagagactactccgctcgttttaagagagacgctgcggcgctc 1745
1811 ccgactcttcgaacacgcgcgtgacagccgtcaacttcgaacgagagaaagccgaggggttc 1870
1746 ccggttgccagagcgtctccgcctcgacagagacatcgacagcgtagtagtgcattcatctg 1805
1871 ttctgcgcgaagggcggtgtgcacacagacagagacatccggggtgtcgtcagacattta 1930
1806 atggggcaacgctgtgtactacgtgttcatagatgaatacaacgggagagactcttgtagc 1865
1931 ttgggttaacgctgttcaacttctctgacacacaaactcaacccgagacatccgcgtctca 1990
1866 atggaacgacttccttcgaacacacgagacagacacactatctgttcaagctgcgcgtac 1925
1991 gctaaacgactcttcacataccatcgagaagaagtgaattatttggtagagtgtagcgtactca 2050
1926 gatactttggcgagacgcccacacaaacagtgacacccagcgtgtgtatgaactgtgag 1985
2051 gacacgctgtggca---accccttccacagcgacgctacagcgtgtcatcagaactaaat 2107
1986 gatgtcaacacacactctccacactacgtcttgcgcgttcgacactcaagcgttcgagag 2045
2108 gacatcaacaacacgcccacacactttagcgtctcgagtcgagtcgacagtggaagag 2167
2046 aacgttcccgagagtagagatatcccgaggaatacactgactgaacccgagacacgac 2105
2168 aacgtgctgtatgacagctcatcaacccaagaggttaacgcccacccgacccgacacg 2227
2106 gctcaacgctgttcgaagatcgactggagactccacactgggcccacaaagcgagagag 2165
2228 gccgcttcgctgtcgatataactggagacactcttcgcgcaacgaagcgccgacag 2287
2166 accaacctactgaatacgttcggtgtatagtatatacgaagaatatacccccacgagggc 2225
2288 gctaaccccgagagtttagaatgtcgltggaatcgaagacacatctcccccagataaac 2347
2226 aacgggggttccgacacggcgctcgltgtgtgcaagagatccggagacaacgtccacac 2285
2348 aacgggggttccgacacggcgctcgltgtgtgcaagagatccggagacaacgtccacac 2407
2286 gacttcgaggaatttcgagatgtcttaactccacgctccggtgagggagactccaacgtgc 2345
2408 gactacggagagtttgaggtctctccctccacagtgaggggtgctgacacttaacgcgtc 2467
2346 atcgagatgaatacgaatgagcggaagtcacagatcaacataatctgacatgtaacgacaa 2405

2468 taaggagacgaactaacgacgaatcgatgtctcaaatataataatccatatgtaacgacaa 2527
2406 ggcgcgaactcttcgaaacggaacgctgacgagacagatgtgcgtgtgcgagcgtgcggcc 2465
2528 ggcgcggt 2587
2466 agcgacgctcatcgctccgctgcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 2525
2588 ggcgggt 2647
2526 gttgcgtacatacaacacataagaacacacacacacacacacacacacacacacacacacac 2585
2648 gtcgcgtac 2707
2586 acaactgtcaaatgtgagtgtagacgagacgagtgtagacgagacgagacgagacgagac 2645
2708 cctcaacggttcaaatcttcgcgt 2767
2646 ttctactgttactac 2705
2768 ttccacacttactatacaagtggt 2827
2706 tgtctctcagaattccaaactacttcgaagttccaggcgaatccgaatacgaatacctcgaac 2765
2828 tgcgcccttcgacccgacttaatttggaacacggaagaaataacacacacacacacacacac 2887
2766 acaacacaaagatgctctgtagacgctcgaactgtagaagtccaacagacggtgtacgtctg 2825
2888 acgacacaaagatgctccgacgagcgaaacgacacacacacacacacacacacacacacac 2947
2826 gagaatgtccagagcgtgtagagagagtggttccagctgtactccacacacacacacacacac 2885
2948 gagaacgac 3007
2886 gattgtac 2945
3008 gaatatatccac 3067
2946 ttcttcgagtgtagac 2996
3068 ttctctctcgtgaacacgagacacgacgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3127
2997 gacgaaatattgacacgagtgtagagtagagcctacacacacacacacacacacacacacac 3056
3128 ggcaggt 3187
3057 gataactcttctcgtatgt 3116
3188 gacaacttcacatgt 3247
3117 gttctatgt 3176
3248 atctgt 3307
3177 gatgttccgaaagagcaggt 3236
3308 actatatacgtgaac 3367
3237 aggtatgaac 3296
3366 cgcgacgagcccggaac 3423
3297 accgacgagacatgtaggtgtgcgagatctctcacacatgattcgtatgtgaacacaaact 3356
3424 ----- 3423
3357 gggtgaactggaacgctatggaactggaagggtatgtggcgacttaacgaatatcatct 3416
3424 -----gcgcgtgaac 3469
3417 gaggccttcgacacgctacacgacgagcagaggttccaaagagacgtataccctgtgtatc 3476
3470 ctggacttcgac 3529

[illegible]

QY	4557	gagacgcttgagcgcgaacacagaatttcataatgcgcgcacgtttatgaacccgggttcaactg	4616
Db	4601	caacacaggttcgacaacacacagagactttatcgtcgacacattcaagcgtgtgttcaactg	4660
QY	4617	acgttcacacatcgaccacaggtgtgtgtccgtttcagaga---cagcaagcggcgttgcgcaaaag	4672
Db	4661	acctgcacaacatcgaccacaggtgtgtgtcccgcttaacagacccttcacggcgtgtgtgtgtgag	4722
QY	4674	gacacacacggaggtgcgcgcgcacattcatccggagacaaacgtgcccgttcgagacaaagag	4733
Db	4721	cacagcagcgcagatg-gcgcgcaccttcataatcgagacaacgtaaccgttactcgtatgag	4779
QY	4734	gtcgaagcgcgtcgcgcgcagacacaggtgtgtgtcgtgcgcacatccagctgtatgtgaacc	4793
Db	4780	atagacagatccgtagtgacctagttctctctgtagcttcatacaacaacagcgtgtgtgc	4838
QY	4794	aacggccgtgtgtgcgcgaagacctgtgtgagcgtgtgacacatccgaactagtcgagagatca	4855
Db	4840	gcgacgcgtgtgtgtgcagagactgtgtgacacatccagccg-----gactgtctg	4890
QY	4854	atgcagatcgcgcgtctcagcactagacacgcgcgcctccgcgtgtgtagcttccctctgcctc	4913
Db	4891	actcgagccttaacggtgtatgcttcgtcctcaactgtctgcgt-ctcggtttcaatgtgcctt	4948
QY	4914	gtactgcttctcgcatgttctgttaagacaagaacgtgaacccgcgcgcgttcgaacatc	4973
Db	4950	gtgctactgtcttaccttcatacatcaagactagacgcgttaacgcagcgtgttgaaagccctg	5005
QY	4974	tccatgcagcaagt-acgcgttcgtgtgacttcggcttcggaaccgcgcgcgcgtc---ggcgc	5028
Db	5010	tcgatgcagcaaggtgaaggtcttcaactgtgacttgcattgtaaccgcgcgcgcgtatgcgcgcc	5065
QY	5030	gggcgcacaagaacgcgcgttcgagggctccaaccccatgttgaaacgaagccatccgcgc	5088
Db	5070	cgcgcacaacaacacacactgttgaaaggtctccaacctatcttcattgaacacaataaagac	5129
QY	5090	gcccgcacttcagcgcacatcagttgacgcgaatgtggcgactccgcacctgatacgtatcgaaga	5149
Db	5130	gccagatttagatgcatctagtcgaggggttcccaacgacatctgatacgtatcgcatacgaaga	5189
QY	5150	catgcc 5155	
Db	5190	tcttcc 5195	
RESULT 6			
AAE30940			
ID	AAE30940 standard; cDNA; 1604 BP.		
XX			
AC	AAE30940;		
XX			
DT	09-JUL-2001 (first entry)		
XX			
DE	Ostrinia nubilalis Bt toxin receptor partial cDNA.		
XX			
KW	European corn borer; ECB; <i>Plutillus thuringiensis</i> ; Bt toxin;		
KW	receptor; crystal protein; CryIa; biological control; insecticide;		
KW	crop protection; ss.		
XX			
OS	Ostrinia nubilalis.		
XX			
PN	MO200136639-A2.		
XX			
PD	25-MAY-2001.		
XX			
PF	17-NOV-2000; 2000MO-US31674.		
XX			
PR	18-NOV-1999; 99US-0166285.		
PR	21-SEP-2000; 2000US-0234099.		
XX			
PA	(PION-) PIONEER HI-BRED INT INC.		
XX			
PI	Flannagan RD, Mathis JP, Meyer TE;		

XX WPI; 2001-329223/34.
XX
XX New nucleic acid encoding Bt toxin receptor, useful for screening and
PT identifying new Bt toxin receptor ligands useful as new insecticidal
PT toxins -
XX
PS Example 1; Page 83; 85pp; English.
XX
CC The present sequence of is that of a partial cDNA clone obtained
CC by RT-PCR amplification of European corn borer (ECB), *Ostrinia*
CC nubilalis, larva midgut cDNA. The clone was used to screen a
CC midgut cDNA to isolate further clones, from which a full-length
CC sequence (see AAF30933) encoding the Bacillus thuringiensis (Bt)
CC toxin receptor (see AAF20498) of ECB was deduced. The polypeptide
CC encoded by the present clone was also used to inoculate rabbits for
CC polyclonal antibody production. The ECB Bt toxin receptor binds
CC CryIA, particularly CryIA(b), toxins. The invention provides
CC compositions and methods for modulating susceptibility of a cell
CC to Bt toxins. The compositions include Bt toxin receptor proteins
CC from the lepidopteran insects ECB, corn earworm (*Heliothis zea*) and
CC all armyworm (Spodoptera frugiperda). The methods are useful for
CC investigating the structure-function relationships of Bt toxin
CC receptors, investigating toxin-receptor interactions, elucidating
CC the mode of action of Bt toxins, screening and identifying novel
CC Bt toxin receptor ligands including novel insecticidal toxins, and
CC designing and developing novel Bt toxin receptor ligands. The
CC methods are useful for managing Bt toxin resistance in plant pests,
CC and protecting plants against damage by plant pests.
XX
XX Sequence 1604 BP; 386 A; 465 C; 419 G; 334 T; 0 other;
SQ
Query Match 28.1%; Score 1547.4; DB 22; Length 1604;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 1566; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
Q 3040 tcttgaattcgtacatactcttcttgatgtagtgcgtgagaaaccagacg 3099
Db 8 tcttctcaacctcagacacactcttcttgacgtgtaggagaaaccagacg 67
QY 3100 aagtgaaatttgtcgtctctatgtgatgtgaacgaacagctctcgtatgcattgc 3159
Db 68 aagtgaaatttgtcgtctctatgtgatgtgaacgaacagctctcgtatgcattgc 127
QY 3160 ctgataactcgtgttatgtttccgaaaggacgtgctggttcgttaccacag 3219
Db 128 ctgataactcgtgttatgtttccgaaaggacgtgctggttcgttaccacag 187
Q 220 aaatctacgcacggacggatgtgaacgaacgaacgaacgtgtgtcggttaaggaa 3279
Db 188 aaatctacgcacggacggatgtgaacgaacgaacgaacgtgtgtcggttaaggaa 247
QY 3280 tctcgtacctacgatacgcacgacgacatcgtggtcggtatctcttcaacatgatct 3339
Db 248 tctcgtacctacgatacgcacgacgacatcgtggtcggtatctcttcaacatgatct 307
QY 3340 cgattgaaacaaactcgggaacttgagacgcgtatgacttggggggtatgggca 3399
Db 308 cgattgaaacaaactcgggaacttgagacgcgtatgacttggggggtatgggca 367
QY 3400 cttaacaaattcattagagccttcgacacggtacccgacagagaggtccaaacgga 3459
Db 368 cttaacaaattcattagagccttcgacacggtacccgacagagaggtccaaacgga 427
QY 3460 cgtacacccgtgcatccgcacctacaaacttccacacccgtgtgtcgttcccgcaac 3519
Db 428 cgtacacccgtgcatccgcacctacaaacttccacacccgtgtgtcgttcccgcaac 487
QY 3520 ccgactccgttactcgtgttcttagggagcgacgaagaaggcggtctctggagacgg 3579
Db 488 ccgactccgttactcgtgttcttagggagcgacgaagaaggcggtctctggagacgg 547

QY 3580 ctgcaacgaattctctgagcgaatctacgacacgaagacggtccctccacggtgca 3639
Db 548 ctgcaacgaattctctgagcgaatctacgacacgaagacggtccctccacggtgca 607
QY 3640 gcttcacgttccacgttccaggaagaatgagagccgttcgacttggatatactgaag 3699
Db 608 gcttcacgttccacgttccaggaagaatgagagccgttcgacttggatatactgaag 667
QY 3700 tgggagcaggaagaatgagagccgttatactacgacacgttctccaggaagaatca 3759
Db 668 tgggagcaggaagaatgagagccgttatactacgacacgttctccaggaagaatca 727
QY 3760 gacattcagatcagatcccgagccacgagcagcagcagcagccgcgtttgga 3819
Db 728 gacattcagatcagatcccgagccacgagcagcagcagcagccgcgtttgga 787
QY 3820 ccgagtcacgttcttggtgtgtcttctgtaaccacagagggagaccgtgttcagcga 3879
Db 788 ccgagtcacgttcttggtgtgtcttctgtaaccacagagggagaccgtgttcagcga 847
QY 3880 atgcagctactgtcgcctctctcagaggtggaagaagcctcgttgaagatttgaagtc 3939
Db 848 atgcagctactgtcgcctctctcagaggtggaagaagcctcgttgaagatttgaagtc 907
QY 3940 cgcagcagaagacgttaaaaaaacacactctgcgaagatgactgccaagatatctactca 3999
Db 908 cgcagcagaagacgttaaaaaaacacactctgcgaagatgactgccaagatatctactca 967
QY 4000 ggttattgagcgaacaaagaggtcttctcgtactgtagcagatcaagacgttactc 4059
Db 968 ggttattgagcgaacaaagaggtcttctcgtactgtagcagatcaagacgttactc 1027
QY 4060 cccttgcgagaggttgcgcgcgaggtgcgcgcgtcttaccagcttcacacgtcgcgcga 4119
Db 1028 cccttgcgagaggttgcgcgcgaggtgcgcgcgtcttaccagcttcacacgtcgcgcga 1087
QY 4120 gaaactgcgcgcgcgcacgtggatccctctgcgaacttccatccctcgttgcaggtca 4179
Db 1088 gaaactgcgcgcgcgcacgtggatccctctgcgaacttccatccctcgttgcaggtca 1147
QY 4180 atgtaagaagaagcgaacccgcgcacatcttcgaagaagacgttcaacacgcggacatt 4239
Db 1148 atgtaagaagaagcgaacccgcgcacatcttcgaagaagacgttcaacacgcggacatt 1207
QY 4240 cgcagcttgacagcattgcccggagatgtctactgttcaaggcgagccacacagaagaagc 4299
Db 1208 cgcagcttgacagcattgcccggagatgtctactgttcaaggcgagccacacagaagaagc 1267
QY 4300 aacacatcagtaacacatagacgttcgacagatgacgttgcagacagcgttagaagccg 4359
Db 1268 aacacatcagtaacacatagacgttcgacagatgacgttgcagacagcgttagaagccg 1327
QY 4360 tgcgcgactcgcgttcgctgcgttcgacacacgcgggtgtcttgcgttaataagagc 4419
Db 1328 tgcgcgactcgcgttcgctgcgttcgacacacgcgggtgtcttgcgttaataagagc 1387
QY 4420 ccacgcgttccatgacagcagatgttcgagttcgaagtcacgtatcgtacagatacagcttg 4479
Db 1388 ccacgcgttccatgacagcagatgttcgagttcgaagtcacgtatcgtacagatacagcttg 1447
QY 4480 caatcagacagcccggtgtgaaggtctactcctatctatcgcgaacaaacgcgttgacattca 4539
Db 1448 caatcagacagcccggtgtgaaggtctactcctatctatcgcgaacaaacgcgttgacattca 1507
QY 4540 ttctcgataacacacttgagacgttgcagcagaagaagaatttcatagcgcgcagcttca 4599
Db 1508 ttctcgataacacacttgagacgttgcagcagaagaagaatttcatagcgcgcagcttca 1567
QY 4600 gcaacgggttcaacatgacgttcgaacatcgaacaggt 4636
Db 1568 gcaacgggttcaacatgacgttcgaacatcgaacaggt 1604


```

RESULT 7
AAF30935
ID AAF30935 standard; cDNA; 5592 BP.
XX
XX AAF30935;
XX
XX 09-JUL-2001 (first entry)
XX
XX Spodoptera frugiperda Bt toxin receptor cDNA.
XX
XX Fall armyworm; FAW; Bacillus thuringiensis; Bt toxin;
XX receptor; crystal protein; CryIA; biological control; insecticide;
XX crop protection; ss.
XX
XX Spodoptera frugiperda.
XX
XX Key Location/Qualifiers
XX CDS 162..5366
XX sig_peptide 162..227 /*tag= a
XX mat_peptide 228..5363 /*tag= b
XX /*tag= c
XX
XX WO200136639-A2.
XX
XX 25-MAY-2001.
XX
XX 17-NOV-2000; 2000WO-US31674.
XX
XX 18-NOV-1999; 99US-0166285.
XX
XX 21-SEP-2000; 2000US-0234099.
XX
XX (PION-) PIONEER HI-BRED INT INC.
XX
XX Flannagan RD, Mathis JP, Meyer TE;
XX
XX WPI: 2001-329223/34.
XX
XX P-PSDB; AAB30500.
XX
XX New nucleic acid encoding Bt toxin receptor, useful for screening and
XX identifying new Bt toxin receptor ligands useful as new insecticidal
XX toxins -
XX
XX Claim 1(a): Page 72-79; 85pp; English.
XX
XX The present sequence of is that of cDNA encoding the Bacillus
XX thuringiensis (Bt) toxin receptor (see AAB20500) of the fall
XX armyworm (FAW), Spodoptera frugiperda. The receptor protein binds
XX CryIA toxins, particularly CryIA(b) toxins. The invention provides
XX compositions and methods for modulating susceptibility of a cell
XX to Bt toxins. The compositions include Bt toxin receptor proteins
XX from the Lepidopteran insects FAW, European corn borer (Ostrinia
XX nubilalis) and European corn borer (Heliothis zea). Nucleic acids
XX encoding the polypeptides, antibodies specific to the polypeptides,
XX and nucleic acid constructs for expressing the polypeptides in cells
XX are also provided. The methods of the invention are useful for
XX investigating the structure-function relationships of Bt toxin
XX receptors, investigating toxin-receptor interactions, elucidating
XX the mode of action of Bt toxins, screening and identifying novel Bt
XX toxin receptor ligands including novel insecticidal toxins, and
XX designing and developing novel Bt toxin receptor ligands. The
XX methods are useful for managing Bt toxin resistance in plant pests,
XX and protecting plants against damage by plant pests.
XX
XX Sequence 5592 BP; 1662 A; 1301 C; 1332 G; 1297 T; 0 other;

```

```

Query Match 26.7%; Score 1465.8; DB 22; Length 5592;
Best Local Similarity 58.0%; Pred. No. 3.7e-312;
Matches 2969; Conservative 0; Mismatches 2037; Indels 111; Gaps 17;
Oy 160 aatgagggttgaagaggtcttcacagcagtgctactgtctcttaagcctctgcgcgac 219

```

```

Db 160 agatggcgttcgatgtgcaatactcagacagcaaatctgtctgtactaccactgtcagc 219
Oy 220 tagccaaccaagatgttctgtacattatcgcgaataccagaccggagactccggactgc 279
Db 220 cacagcgagatcgtatgtgtctacatcgtlagaataccagaccagagcgtctactcc 279
Oy 280 cgcctatgttaacgaagaataatcagtagtgaacagctcttaataccgcccggacc 339
Db 280 caactcaaaatttgcaggtttaaactgtggtccagagccactattaccagctgagatc 339
Oy 340 gagaggaatglatgtatggaaactctctacggatcaaatga-----ttc 384
Db 340 gagaagaggttcgcctccaatgtactatgaaccctgataccctggagcaacaacatgtgacc 399
Oy 385 aggtcatatcatatgaggaagaagaatcgaagagacgtcatcttcggaagcttaactatc 444
Db 400 agagaatttacaatgagagagatcgaaggtcccgtaacttgcgaataattaactacc 459
Oy 445 aagggtccaacacgcgcggtgtcgtcgtatgtatgaagccagccagagccca---gtctg 501
Db 460 aaggaaacacccctcctcaaaataagattaccttctgtgtgtgtaagccacatgtctg 519
Oy 502 gccctgagttcgc---acagaatgaagcagacggccaatgagccttgttatacga 558
Db 520 gagcagaatattcgtgaataatccttgacgcgaactgagactgtgatttgaataactca 579
Oy 559 gacaagactacgagacgaacgaacatcgcagagctatgttctcaatccaagtsgaggtg 618
Db 580 ggcagagcatatgaactcctgtatgcagagatacacgttctgagtgtgaggaagcc 639
Oy 619 aatcagagcgtactgtgtggtggtgagatgtcaaatcgaacgaacatccgcccactc 678
Db 640 agtcctgtgttgaacggtgagagctggatgtgtgaacatcgcagcaaatcgcgcactca 699
Oy 679 tgaaggtgttcagcgctcgtatccgaacatcgcagagcgtgactgacgcagactgcg 738
Db 700 ttgagatgtttgaagccttgcgaactacccggaactgttgaacccatgtttaagaatgta 759
Oy 739 tgtaccaagtgtcagaccgcgcagcgtgtgaatcagacaccgccttcatgacgtctgc 798
Db 760 aatatacgtgtccgcagcgaacggttcgtatcatcaagtgttatgtatcatatag 819
Oy 799 acagcagacagcgttcagatgaagaacatcttcatatggttggagaatacgaacccagc 858
Db 820 acagcagagag---aggaaacgaagaagaatctcgaactgtacgaagaagattaccggcg 876
Oy 859 actgttcaatatgaagaatgactgtgggatatcattcgccttgaactcagacgaactc 918
Db 877 attgacgaaggtgtatattgtcttgattgaatgaanaaatcctcttgattagaagaatc 936
Oy 919 agcttcataatttagcgtcacacgcttctgactcgtctacgaacaacacacagctacaa 978
Db 937 ctctacacatttcagaaatcaaggtcttgcattccttcaacaaacatgagacgtgtca 996
Oy 979 tgatgtgcaagtgaagaacgtagaggtctcgcccccgcgcgggttgtagatctctcag 1038
Db 997 tgaatgttgaagtaagaagaacgtagaataatcctccgtgtgtagatcttctgtcgt 1056
Oy 1039 tgcacagctttgacagagaactaatacagagcttctcctccgcgcgtagaagggagca 1098
Db 1057 tgcacagctttgtagtaaaacagcgggaatcgttccacagctgcgagctattgtatgcgca 1116
Oy 1099 cgggaatacaatagggcatcaactataccctcaatcagagatgaagctcagcaactcttt 1158
Db 1117 cgggaatacaatcaactatctatcgtlatagaactgaagatgaagaacaaagatctct 1176
Oy 1159 -----ccctgaggtgtatgaagcagagcattctctgacgttgactgacgtc 1206
Db 1177 tcaagcattgagaacacataggggaagcagagacggtgcagacattccaagctgcctatag 1236
Oy 1207 accgcgaacagcttgaagagacgttttcaacccctacacatcgttgc----- 1252

```

Dh	1237	acgagactacccgnaaaggatattgtttccatataagaataattgcatataacaaggtg	1296
Qy	1253	---ttacaatctactgacgctagctttgtcaacagagggccacatttccatcgtca	1308
Dh	1297	ataatgccaagaaggtgaattcatcgttcgagacctcagccaatgtgacgtttataat	1356
Qy	1309	acgagctcaattgatcagcgaccggagcgctgcataaagaataacgatttgaatatcgtg	1368
Dh	1357	acgatacaaatgtatcagagggccgaacaccttccataaagaatacagatctcccaatg	1416
Qy	1369	aggaatctccaatgactctaaacttccaaatctgaagaatttggatttccatgatcgaattgg	1428
Dh	1417	aagaactcgtgactgaccttgaattg---caagagtttggtttccatgacccgtgaacttg	1473
Qy	1429	gtgaaaacgcttcaataacacagtgaagaacttgaaggacgtgttcccgcaaggggcggtccg	1488
Dh	1474	gtccccaacgctcagtaacgacttccatttagagatatacagccagaggggcccatacgg	1533
Qy	1489	cattctacatcgcgcggggagcggtctaccagaagcgagaccttccatctggtgcacataa	1548
Dh	1534	ctttctacatcgcgccttgaaagaggttaccagggcccgatcttccacataagtaactga	1593
Qy	1549	accacacatcgtcgtgaattgaagaagtgtacatt---ttcagaacatccatgaatga	1605
Dh	1594	tccaatacactgttggtattatgaagaatggaagacttacagacaggaataaagctaaaggcag	1653
Qy	1606	aagcagtgagacatgaacaacgcttagccacagctgggcgagggcgctgtgtgacgtgaacctga	1665
Dh	1654	tagcaattgacagacagacataacatacactatgtgggaagcaatttaacattaccta	1713
Qy	1666	tcaacttggaaacgacgaacttccacttccgagagagacagctaccgcgctcogtttaagg	1725
Dh	1714	tcaattggaatgtagagctaccataattccgagagagacgcttacaacagctgacatttga	1773
Qy	1726	agaccgctcgcgccgggttcctccgggtggccaagctgtgctgcctccgaacgaacattcgacg	1785
Dh	1774	agacggtccgggtgagtgcttccacatttggtaataacggggcttaaaacagagacacgggtg	1833
Qy	1786	acgtgagtagtgcattctcatatgtgagcaacgctgtgtgactaccgtgttcaatagatca	1845
Dh	1834	acatagtcggagcactcgtatatttgggcaacgctgcaaaccttccgtggaattgacatagata	1893
Qy	1846	cgggagagacttcgtgtgagcatgagacgagtcccttcgactaccacagacagaacactctat	1905
Dh	1894	ctggagagtggtgacgtgtcagcgagcacttacttattatataaagacagaaagaaatca	1953
Qy	1906	tgtgttaaggtgagcgctcgagacatactttgggagcgagcgcccaacaacaacagtgacaccc	1965
Dh	1954	taagtctagatctcgtgcgtgtgtgatacacttag---ttacttaagaacaggggtctacacac	2010
Qy	1966	agctcgtgataagactcgtgagtagtgcataacaacactctccacaccttaacgttgcggcgtt	2025
Dh	2011	agctcagagataatttttggaaagataataacaacacgcccacactctactcgtgactgtccaogtt	2070
Qy	2026	cgactccaaagcgttcgagagagaacggttcccgaaagagatacagatatccgggaaatcacctg	2085
Dh	2071	ccagttccaaagtgtagaagagaacggttgaagctcgggagcccgaattacgaggggtctaaag	2130
Qy	2086	ctaccagcccggaacacacagcgcttaccigtgtgttgagatgacaggggaattccacactg	2145
Dh	2131	cgacagaccccgaaataccacacagcgacttccacttcagctgagatctgattgggaacattcttcag	2190
Qy	2146	ccaccaaagcaag---cagagagacacacacctactcaatagctcggtgtgatagtatcgt	2202
Dh	2191	ctacgaagccagggacccaatgtagcccaacactcgcagactaccacaggaatgctgtagaatcc	2250
Qy	2203	aaacgataataccccaacggaaggaacacgggtgtccgccaatcggtgcctcgtgtgtgacaag	2262
Dh	2251	tgacggtatacccgactcgtcgtacaatacagggagctgaggtgactgtgtgcaagctg	2310
Qy	2263	agatccgggacaacgctcaacatacgtactcggaggaattcgaagatgcttaccatacgcgtcc	2322
Dh	2311	aggtcgaatgtgaggtgtacatctcgattacagagaaggttttgaagttgtcgtacatcgtctgcca	2370

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2002, 13:48:22 ; Search time 74.72 Seconds

(without alignments)
1702.139 Million cell updates/sec

Title: US-09-715-909-2

Perfect score: 8909
Sequence: I-MGVERFPAVLVLSASAL.....TGNNGFGKSTPYLPQHPK 1717

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

A.Geneseq-1101.*
1: /SIDS2/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8909	100.0	1717	22	AAAB20498
2	5179.5	88.9	1730	22	AAAB20499
3	4923	55.3	1729	22	AAAB2331
4	4693	52.7	1528	20	AAAB20182
5	4561.5	51.7	1734	22	AAAB20500
6	4408	49.5	1528	17	AAAB5333
7	519	5.8	2923	22	AAU07054
8	519	5.8	2956	22	AAU07053
9	494.5	5.6	3298	22	AAE03657
10	459.5	5.2	2707	18	AAW27161
11	444.5	5.0	3014	22	AAU02196

12	324	3.6	1026	15	AAAB58906	Human protocadherin
13	324	3.6	1026	17	AAAB7146	Protocadherin clon
14	324	3.6	1203	15	AAAB58911	Product of alterna
15	324	3.6	1203	17	AAAB7152	Alternatively spli
16	317	3.5	2405	21	AAAB2192	Human ORFX ORF1956
17	311	3.5	1180	19	AAAB61273	Homo sapiens proto
18	305	3.4	788	22	AAAB48370	Human SEC3 protein
19	303	3.4	916	20	AAAY41743	Human PRO707 prote
20	303	3.4	916	21	AAAB4299	Human PRO707 (UNO3
21	300	3.4	627	21	AAAB38349	Human secreted pro
22	296	3.3	896	19	AAW41252	Xenopus paraxial p
23	295.5	3.3	797	17	AAAB6866	Human protocadheri
24	291.5	3.3	882	20	AAAY09375	Wild-type E-cadher
25	291.5	3.3	882	21	AAAB35730	Human E-cadherin a
26	291.5	3.3	882	22	AAAB73490	Human E-cadherin
27	288	3.2	100	22	AAAB82332	Pink bollworm toxi
28	285	3.2	558	22	AAAB38969	Human polyepitide
29	284	3.2	772	20	AAAY41725	Human PRO941 prote
30	284	3.2	772	21	AAAB44281	Human PRO941 (UNO4
31	284	3.2	772	22	AAAB31209	Amino acid sequenc
32	283.5	3.2	889	20	AAAY24913	Human ontherin. H
33	283.5	3.2	889	20	AAAY21687	Cadherin-like poly
34	282.5	3.2	1184	20	AAAY41750	Human PRO731 prote
35	282.5	3.2	1184	21	AAAB44306	Human PRO731 (UNO3
36	282.5	3.2	1184	22	AAU12368	Human PRO731 polyp
37	282	3.2	1187	22	AAE06028	Human asthma-assoc
38	282	3.2	1187	22	AAAB95684	Human protein sequ
39	280	3.1	831	22	AAAB31106	Peptide #5143 enco
40	277	3.1	800	21	AAAB18920	A novel polyepitid
41	274.5	3.1	653	18	AAAB25657	Rat cadherin-10.
42	274.5	3.1	653	18	AAAB13128	Full length rat ca
43	274.5	3.1	878	15	AAAB55060	Sequence of human
44	274.5	3.1	878	16	AAAB5487	Human E-cadherin p
45	274	3.1	906	21	AAAY0741	Human N-cadherin.

ALIGNMENTS

RESULT 1	AAAB20498	standard; Protein: 1717 AA.
ID	AAAB20498;	
AC	09-JUL-2001	(first entry)
DT	Ostrinia nubilalis	Bt toxin receptor.
DE	European corn borer; ECB; Bacillus thuringiensis; Bt toxin;	
KW	receptor; crystal protein; CryIa; biological control; insecticide;	
KW	crop protection.	
XX	Ostrinia nubilalis.	
OS	Key	Location/Qualifiers
XX	Binding-site	1293..1462
FT	Domain	/label=CryIa_binding_site
FT		1571..1589
FT		/label=Transmembrane_domain
XX	MO200136639-A2.	
XX	25-MAY-2001.	
FD	17-NOV-2000; 2000WO-US31674.	
XX	18-NOV-1999; 99US-0166285.	
PF	21-SEP-2000; 2000US-0234099.	
XX	(PION-) PIONEER HI-BRED INT INC.	
XX	Flannagan RD, Mathis JP, Meyer TE;	

XX MPI: 2001-329223/34.
DR N-PSDB: AAF30933.
PT New nucleic acid encoding Bt toxin receptor, useful for screening and
PT identifying new Bt toxin receptor ligands useful as new insecticidal
toxins.
XX
PS Claim 4(a); Page 57-61; 85pp; English.
CC The present sequence of is that of the *Bacillus thuringiensis* (Bt)
CC toxin receptor of the European corn borer (ECB), *Ostrinia nubilalis*.
CC The sequence was deduced from cDNA isolated from an ECB larva
CC midgut cDNA. The receptor protein binds CryIA toxins, particularly
CC CryIA(b) toxins. The invention provides compositions and methods
CC for modulating the susceptibility of a cell to Bt toxins. The
CC compositions include Bt toxin receptor proteins from the
CC lepidopteran insects ECB, corn earworm (*Heliothis zea*) and fall
CC armyworm (*Spodoptera frugiperda*). Nucleic acids encoding the
CC polypeptides, antibodies specific to the polypeptides, and nucleic
CC acid constructs for expressing the polypeptides in cells are also
CC provided. The methods of the invention are useful for investigating
CC the structure-function relationships of Bt toxin receptors,
CC investigating toxin-receptor interactions, elucidating the mode of
CC action of Bt toxins, screening and identifying novel Bt toxin
CC receptor ligands including novel insecticidal toxins, and designing
CC and developing novel Bt toxin receptor ligands. The methods are
CC useful for managing Bt toxin resistance in plant pests, and
CC protecting plants against damage by plant pests.
XX
XX Sequence 1717 AA:
S0
Query Match 100.0%; Score 8909; DB 22; Length 1717;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1717; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGVERFPAVLVLSAALANORCSYIIAPRPETPLPIDYEGKWSQPLIGPTR 60
DB 1 mgverfpaavlvsalasaaanqrcsyiaiprpelplidyegkwsqpligptr 60
QY 61 EEVCKENLPPQOMIOYVMEETBEDVIAKLNQSGMTPLVSTMSGPRQAGFEFQON 120
DB 61 eevckenlppqomioyvmeebedvialklnygsnltplvstmsgprqagfefqon 120
QY 61 EEVCKENLPPQOMIOYVMEETBEDVIAKLNQSGMTPLVSTMSGPRQAGFEFQON 120
DB 61 eevckenlppqomioyvmeebedvialklnygsnltplvstmsgprqagfefqon 120
QY 121 EADGWSLVITROODYEFATQOSYVESIOVEGSAVLVAETIVINDNPILQVVSACV 180
DB 121 eadgwslvitroodyefatqosyvesiovegavlaetivindnpllvvsacv 180
QY 181 IPEHGEARLITDCVYVSDRDGEISTRFMTFRVDSGRADESIFYMVGEYDPSDMFMKMT 240
DB 181 ipehgearlitdcvyvsdrdgeistrfmtfrvdsgradesifymvgeydpdmfmkmt 240
QY 241 VGINSPLNFETTQHLIFSVTASDLSLPNNHTVTMVOVENESRPPRWYEIFSVQFDEKT 300
DB 241 vginsplnfettqhlifsvtasdlsppnnhtvtmvovenesrpprwyefsvqfdekt 300
QY 241 VGINSPLNFETTQHLIFSVTASDLSLPNNHTVTMVOVENESRPPRWYEIFSVQFDEKT 300
DB 241 vginsplnfettqhlifsvtasdlsppnnhtvtmvovenesrpprwyefsvqfdekt 300
QY 301 NQSESLRAIDGTTGIRNAINVTILRIDDADEFSLVIEDGALHTETIDROKLEBELFNL 360
DB 301 nqseslraidgttgirnaivntilriddaderfslviedgalhtetidroklebelfnl 360
QY 361 TIVAAKSTDPASATPAHTFIIVNDVNDORPEPLHREYSIDIMEETPMTLNFEDEGFHDR 420
DB 361 tivaakstdpasatpahrtfiivndvndorpeplhreyssidimeetpmtlnfedegfhdr 420
QY 421 DLGENAQYTVLEEDVFPFGAASAFYIAPGSGYQROTFLMGITNHTMLDYEDVIFQNIILK 480
DB 421 dlgenaqytleedvfpfgaasafyiapgsgyqrotflmgitnhtmldyedvifqniilk 480
QY 481 VAAVDMNNAASHGEALVYNNLNNMDELPIFEBSYSASFKEFTVAGCFVAVYALMDRI 540
DB 481 vavdmnnaashgealvynnlmmndelpifebsysasfkeftvagcfvavyalmdri 540
QY 481 VAAVDMNNAASHGEALVYNNLNNMDELPIFEBSYSASFKEFTVAGCFVAVYALMDRI 540
DB 481 vavdmnnaashgealvynnlmmndelpifebsysasfkeftvagcfvavyalmdri 540

QY 541 DRYVYHSLMGNAVDYLFIDESTGEIFYSMDADFYHRONTLFFVOVARDPDLGDPHNTVT 600
DB 541 drvyvshlmgnavdylfidestgeifysmdadfyrhontlffvovardpdlgdphtntvt 600
QY 601 TOLVIELEDVNNTPPTLRLPRSTPSVEENVPEGEYSIREITATDPDTSAVLMWEIDMDST 660
DB 601 tolvieledvnnptlrlprstpsveenvpegeysireitattpdttsavlmweidmdst 660
QY 661 WATKQRETNTEVSGIVETIETIYPTEGNRGSAIGRLVQVEIRDNTIIDPEEFMLYLV 720
DB 661 watkqretntevsgivetietyptegnrsgaigrvlvqveirdntiidpeefmlylv 720
QY 721 RYRLNVIQDDYEAETFTTIIDNDNAPFANGTLTQPMRARELAASGLISVYATD 780
DB 721 ryrlnviqddyeaetfttiidndnapfangtltpmrarelaasglisvyatd 780
QY 781 IDGPLYNGVRYLTIPRNTEPGLKIDFTTGQLEVDNNEAIDAEPMRFYLYTVIASDE 840
DB 781 idgplyngvryltiprntepglkidfttgqlvdenneaideapmrfylytviasde 840
QY 841 CSELENFTCEPDSNFEYVPGDIEIEIIDTNKKVPELITEKNTTVYWEANATSGDEVVOL 900
DB 841 cselelntfcepdsnfeypgdieieiidtnkkvpeleteknttvyyweanatsgdevvol 900
QY 901 YSHDRDELVHTVRYTMNFAVNPRLDFEVDLDTGRLLEVHYGDEKLRDGEPTHTI 960
DB 901 yshdrdelvhtvrytmnfaavnprldfefdldtgrllehygdekldrdegepthti 960
QY 961 FVNFIDNFFSDGDRNRDDEVEIEFVLLVDNDNAPEMLPDELRFVDEGAVGVRLPE 1020
DB 961 fvnfidnffsdgdrnrddveiefvllvndnapemlpdelrfvdegavgvrlpe 1020
QY 1021 IYAPDRDEPTDNRKVGIGLIDLTITTDIDIEVPLFMISTENKTGELETFAMDLRGVWGT 1080
DB 1021 iyapdrdeptdnrkvgiglidltittidievpfmlstentkgeletfamdrgvwgt 1080
QY 1081 YEIIFEAFDHGYPOORSNETVTVIRPNFHPVFPQPSVIRLSERATEGGLATA 1140
DB 1081 yeiifeafdhgypoorsnetvtyirpnfhpvfpqpsvirlserateggllata 1140
QY 1141 ANEFLEPIYATDEDEGLHAGSYTFHVQNEEAVQYFDITEVAGENSQGLILRQLEPQIR 1200
DB 1141 aneflepiyatdedeglhagsytfhvqneeavqyfditevagensqglilrqlepeqir 1200
QY 1201 QFRITIRATDGTGTEPGLMTDVTESVVFVPIQGPVSEBNATVAFEGEGELRSFELP 1260
DB 1201 qfriritratdgtgtepglmtdvtesvvfvpiqgpvsebnatvafegegelrsfelpl 1260
QY 1261 QAEDELKHLCEBDQCDDIYRRFTDGNNEGFLVLDSSNYSIQAQELDRVANSYTLHIAAS 1320
DB 1261 qaedelkhlcebdqcddiyyrrftdgnnegflvldssnysiaaqeldrvansytlhiaas 1320
QY 1321 NSPDATGPILOTSLIWMVWVNEANRPRIFEQDLTYAGISTDUSIGRELLTVRASHTEDD 1380
DB 1321 nspdattgpiplotslwvwnvneanrprifefqdltyagistdusigrelltvrashtedd 1380
QY 1381 TITTYTIDRASMOIDSSLEAVRDSAPLHATTGVLSLNNQPTASMHGMEFVDYIATDTASA 1440
DB 1381 tittytidrasmoidssleavrdsaplhattgvlslnnqptasmhgmevdyiatdtasa 1440
QY 1441 IDPARVRYVLISSQNRVFTIPNQLETVYQONNFIAATFSTCFNMTCNDQVVPSSDSG 1500
DB 1441 idparvryvliissqnrvtipnqlenvyqonnfiaatfstcfnmtdcnqdvvpssdsdg 1500
QY 1501 VADOTTEVRAHFRIDBNVVOAEVEAVSDVLLRTIOLMSTNSLVLODVLVTDTPTL 1560
DB 1501 vadottevrahfridbnvvoaeveavsdvllrtiolmstnslvlodvlvtdtpntl 1560
QY 1561 GEESMOIAVYALAAVAGFLCLVLLALFCRTRALNBOALSMTKYGSVDSGLNRAG 1620
DB 1561 geesmoiaavyalaaavgflclvllalfcrtalnbolasmtkygsvdsglnrag 1620
QY 1621 LAPGINKHAVEGSSNPMWNEAIRAPFDALSDASGSDILGIDEMQOFRRDYPFPDDTDS 1680

Db 1621 lappgcnkhavgesnmpmncatirapddaisdsgsdilgiedmpgfrddyirppgdtass 1680
 Qy 1681 SGIVLHMGATDNKPVYTHGNNFGKSTPYLPQPHRK 1717
 1681 sgivlhmgatdnkpvtythgnnfgkstpypilpqpnpk 1717

RESULT 2

AAB20499 standard; Protein; 1730 AA.

AC AAB20499;

DT 09-JUL-2001 (first entry)

DE Corn earworm Bt toxin receptor.

KM Corn earworm; CEM; Bacillus thuringiensis; Bt toxin;

KM Receptor; crystal protein; CryIA; biological control; insecticide;

XX Heliothis zea.

XX Key Location/Qualifiers

XX Binding-site 1305..1471

XX Domain /label=CryIA_binding_site

XX /label=Transmembrane_domain

XX MO20136639-AZ.

XX 25-MAY-2001.

XX 17-NOV-2000; 2000WO-US31674.

XX 18-NOV-1999; 99US-0166285.

XX 21-SEP-2000; 2000US-0234099.

XX (PION- PIONEER HI-BRED-TM INC.

XX Flannagan RD, Mathis, MO, Meyer TE;

XX MP1-2001-329223/34.

XX N-PSDB; AAF30934.

XX New nucleic acid encoding Bt toxin receptor, useful for screening and

XX identifying new Bt toxin receptor ligands useful as new insecticidal

XX toxins -

XX Claim 4(a); Page 68-72; 85pp; English.

XX The present sequence of is that of the Bacillus thuringiensis (Bt)

XX toxin receptor of the corn earworm (CEW), Heliothis zea, which

XX binds CryIA, particularly CryIA(b), toxins. The invention provides

XX compositions and methods for modulating the susceptibility of a

XX cell to Bt toxins. The compositions include Bt toxin receptor

XX proteins from the lepidopteran insects CEM, European corn borer

XX (Ostrinia nubilalis) and fall armyworm (Spodoptera frugiperda).

XX Nucleic acids encoding the polypeptides, antibodies specific to the

XX polypeptides, and nucleic acid constructs for expressing the

XX invention are useful for investigating the structure-function

Query Match 58.1%; Score 5179.5; DB 22; Length 1730;
 Best Local Similarity 59.5%; Pred. No. 0;
 Matches 1033; Conservative 240; Mismatches 426; Indels 37; Gaps 19;

Qy 1 MGVE-RFPFAVLVLSAALANORCSYIATRPETPELPIDYEGKSWKQPLIPPT 59
 1 mavdvrlitaavfliaahlta-qdcsymvaiprperpofpsintdglpwsrlylpvreg 59
 Db 1 REEVCMEFLPDOM--IOYIMEEIEGDDVITAKLNYGSSNP-VLSIMSGPRAQLGPE 116
 60 redvcmeifqpalnprvtvifmeeelegdvalarinygtntplvpsfsgftmlgpr 119
 Qy 117 FRQ-NEADGQMSLVITQRQDYEATMOSYVSIQVGSQAVLALVLEYNIDNPPIQV 175
 120 irrpeengdwhlvitqyqdyetpmpqyifdvrdeplvatvmlivnddnpiilqm 179
 Qy 176 VSACVPIPEHGEARLFDYGVYOVSDRGEISTRMFRVDSRAADSIFFVMGEYPSDMF 235
 180 fepcdipergelglsckylvdsadgelstrmrfeissdtdde-yfeivreniqgwm 238
 Qy 236 NMKMTVGINSPLNFETTLHFSVTSASDLPNNHTVTMMVOYENESRPPRWELFSYQO 295
 239 yhmrvhvkpdyeenplhlfrvayalslpthtvmtvqvenenprrwelfavqg 298
 Qy 236 FDEKTNOSFSLRAIDGDTGINRAIVTLIRDA--DFFSLYEIE--DGAILHVTLEIRD 351
 299 fdektnegsfvraidgdgldkpfyrltekegedlfsiqtleqregawfnvapird 358
 Qy 352 KLERELFNLTIVAYKSTP-----ASFATPEAHIFITVNDVNOQREPLHKESTIDMET 405
 359 tlekevinvslaykygnddvessfskcdvllivndvndqaplpreefstaimet 418
 Qy 406 PWTLNFEFGFHDRLDENAQYVLEEDVPPGAASAFYIAPGSGYQROTFTMGFINHT 465
 419 amtlnl-edfgfhdrlidphagytvheslhprrahesfyiapvgyqsgfmgtpnh 477
 Qy 466 MIDYEDVIFQNIIRKAVDANNASHGEALVYVNLIMNDELFFEESSISASKEITVG 525
 478 mldfvepfgnqlqiravaldmdpdkwyaialnklmndelpmfeedvqvsfideleg 537
 Qy 526 AGFPVATVLAIRDIDVYVSHMGNAVDYLFIDESTGEIFVSMDDADYRONTLPQV 585
 538 agfvalvavakrdvqdvvehsimgnavsylvridkegeelfvtenaefnryqnealvq 597
 Qy 586 RADDTLGDGPHNTVTLQVLEEDVNNTPPLRLRPRSTPSVEENVEGYSREITATDP 645
 598 paddtlge-pyntntqtlvklrdintpplrlpratspsveenvpddvlpqtqhatdp 656
 Qy 646 DTSATVMEIWDSTWATKQRETNPTVGVCIYETIYPTBGNRGAIGRLVQOETIDN 705
 657 dtlaelfeidwqnsyaskgrntsketysgcletetlyphnldngnaigvrvrelidg 716
 Qy 706 VTIDPEEFEMLYIATRVYVNDLVMTVIGDDVDEATFTTITIDMDNAPFANGTLTOMRARE 765
 717 vtiqtemfevlylvtivvndltvgedndistftitidmdnppilwewegltqefvire 776
 Qy 766 LAASCTIGSVLAINDIDPLVYNQVYTTIQPRNNTPEGLVKIDFTQTEVDANEADAIDE 825
 777 vaasgvvlgsvlaiddgplvnyvrytltplrdpedlvidfntqsgsvklhgaide 836
 Qy 826 PWRFLYTVTVADESCELENTECPDSNFEVVEDIETELITIDNNKYPPELTKEFNITV 885
 837 pprqnllyvtvasdkcdllvtceppdptyfepgelfthtdcnkkyqvdeakfeatv 896
 Qy 886 YVMEVATGSDVVOYLSDHSDRDELHYRTVMFAVPRLRDFEVDLDGRLEVHPG 945
 897 yivgaddgqnvqlyasaeldrdeiyhkvysynginalnsirldfemldesglvymnta 956
 Qy 946 DEKLDRDGEPTHTLVFNFDNFSDGGRNDEVEILFVLLDVNDNAPMPLPDELRF 1005
 957 gelldrdgdepthtlfnvfnidnfygsgdnrnetqylvllidndnype-1petlpw 1014

FT Peptide 697..705 specifically claimed in claim 22"
 FT /note= "unique peptide useful as antigen,
 FT specifically claimed in claim 22"
 FT Peptide 886..895
 FT /note= "unique peptide useful as antigen,
 FT specifically claimed in claim 22"
 FT Peptide 1055..1066
 FT /note= "unique peptide useful as antigen,
 FT specifically claimed in claim 22"
 FT Peptide 1321..1331
 FT /note= "unique peptide useful as antigen,
 FT specifically claimed in claim 22"
 FT Peptide 1451..1461
 FT /note= "unique peptide useful as antigen,
 FT specifically claimed in claim 22"
 FT Peptide 1516..1525
 FT /note= "unique peptide useful as antigen,
 FT specifically claimed in claim 22"
 FT Peptide 1572..1582
 FT /note= "unique peptide useful as antigen,
 FT specifically claimed in claim 22"
 FT Peptide 1677..1729
 FT /note= "unique peptide useful as antigen,
 FT specifically claimed in claim 22"
 FT Peptide 291..304
 FT /note= "unique peptide useful as antigen,
 FT specifically claimed in claim 22"
 FT Peptide 622..632
 FT /note= "common peptide useful as antigen,
 FT specifically claimed in claim 22"
 FT Peptide 791..803
 FT /note= "common peptide useful as antigen,
 FT specifically claimed in claim 22"
 FT Peptide 1621..1642
 FT /note= "common peptide useful as antigen,
 FT specifically claimed in claim 22"
 FT WO200134807-A2.
 PD 17-MAY-2001.
 PF 25-OCT-2000; 2000WO-US41521.
 PR 26-OCT-1999; 99US-0161564.
 PR 24-OCT-2000; 2000US-0161564.
 XX (BULL/) BULLA L. A.
 XX (CAND/) CANDAS M.
 PI Bulla LA, Candass M;
 DR WPI, 2001-335932/35.
 DR N-PSDB; AAF30975.
 XX Novel *Bacillus thuringiensis* toxin receptor from pink bollworm, for
 PT identification and design of toxin receptors, for use in combating
 PT emergence of toxin resistance and producing transgenic organisms -
 XX Claim 21; Fig 2A-B; 56pp; English.
 PS The present sequence is that of a novel 194 kDa high affinity
 CC receptor, termed BT-R2, of *Bacillus thuringiensis* (Bt) crystal
 CC toxin CryIAC. The protein is produced in the midgut of pink
 CC bollworm (*Pectinophora gossypiella*), the most destructive pest of
 CC cotton worldwide. The minimum binding fragment (MBF) of the toxin
 CC binding domain on the receptor has been identified. Purified
 CC peptides comprising unique or common motifs of BT-R2 are claimed.
 CC The identification of BT-R2 and its MBF permits the analysis of
 CC receptors in pink bollworm and other insects that affect crop
 CC growth and development, as well as the development of assays for
 CC the cytotoxicity and binding affinity of potential pesticides.
 CC BT-R2 polynucleotides provide the means of manipulating natural

CC and/or introduced homologous receptors, and hence the specific
 CC destruction of organisms, tissues and/or cells of a target
 CC host, including insects resistant to Bt toxins. Transgenic
 CC organisms expressing toxin receptors can also be generated. The
 CC BT-R2 cDNA sequences will also facilitate understanding of the
 CC molecular biology of the toxin receptors in the pink bollworm and
 CC the engineering of more effective toxins in terms of longer
 CC persistence in the field and higher toxicity.
 XX Sequence 1729 AA
 XX
 SQ
 Query Match 56-23; Score 4923; DB 22; Length 1729;
 Best Local Similarity 58.73; Pred. No. 0;
 Matches 986; Conservaltive 239; Mismatches 422; Indels 32; Gaps 18;
 11 LTVSLASALANQ---RCSYII-AIPR-PENPELPIDEGKSWSEOPILGPTRREVC 64
 11 vltfatvfgqetssrcymtdaiprepkddlpdlwsg-gwtdwpllpaeprddvc 69
 65 MENFLPDQMIO-----VIYMEBIEGVYIAKLNYQGSNP--VLSIMSGOPRAQLCPE 116
 70 lngwyp-qltstslgtlthmeeelegdvataklngydsqtpelvgpmwigsnnl-lspe 127
 117 FRQNEADGQMSLVITQRODYEATMQSVFSLQVGEEOAVLVALEIYNIDNPILQVY 176
 128 lr-ne-ngawwyltlnrdyelpmtrtylfdyrpdetrtaarvsistenddnplvrl 185
 177 SACYIPEHGEARLDCVYQVSDRDEISTREMTFVRVDSRADESIFYMVGEDPSDFN 236
 166 dacyfpelgepltcyqvsdsgdgrlsleptfltsdr-edvqlfvyepahlgdfrn 244
 237 MKMTVGINSPLEFETQAHIESVTASDSLNNHTVTMMVOEVNESRPPRWEIERSVOQF 296
 245 mgtlgtltsalnfnesphlfgltaldawpnhvtvmvqyvehrpwmelfavqgf 304
 297 DEKTNQSLRAIDBDTGINRAINVTILIRDDADDFESLEVI--BDGALHMTVEIDRKL 353
 305 demteqgfvraidgdgfygkaihytletdeedlffleclpgghdgalfstamdvdrll 364
 354 ERELFLNTIVAYKSTDAFSAEAFHFIIVNDVQDRPELKHESYIDMEETPMTLNFNE 413
 355 rrdvfrlslvaykydnvafaptrpvlvndlnokkppldgelysismetstpslnfae 424
 414 EGFHDLRLGENAQYTVLEEDVFPFGAASAFYIAFGSGYQKOTFMGTINMTMDYEVY 473
 425 lfgfydeol-lyagelvgqenppveqafyiaptagfqngftaigqdmldyedvp 483
 474 FQNTIIRKAVDMNNAHVGEALVYVNIINNDELPIPEESYSASFKEYGAGFPVATV 533
 484 fgnlkixvatdrdntctfyvaevnnlmmdeepilfeedqlvkkfketvpkdyhvgrl 543
 534 LALDRIDIDVVVHSLMGNAVLYLFIDESTGEIFVSMDDAFYHRONTLFVORADDTLGD 593
 544 rahndldgsvshllgnantflrdeetgdlvaiddafyhqnefnhlyraqdtnse 603
 554 -GPHNTVTQVLEEDVNNTPITLPRSTPSVEENPEGEYSREITATPDPTSAVLM 652
 604 pesrntaaagvlveledvnnptlrlprpsveanpegfehnrelatcdpttaylq 663
 653 FEIMDSTWATKOGRETPTEYVGCIVETIETIPTEGNSGASIGRYVQVEIDNTDIEE 712
 664 feildwtsfackqgdndpiefhgcvdieltfipadrtreavgvvaqglthnvlhnee 723
 713 FEMLYTRVVDLNTVTDIDDEATFTTITIDMDNAPFANGTLQTMRYRELAASGTL 772
 724 feilyltrvrdlhtddgrdydestftvilldmndmpwsgflntqfstrsssgyv 783
 773 IGSVLANDIDGPLYNQVRYTTPRNRPBGLVKIDFTTGQIEVDANEADIDAEWRPRLY 832
 784 lgsvatldidgplynqvrytllpdepgvlgvlfvqgtvdeangaididiprvhln 843
 833 YTVASDSCSLENTRECPDPSNFEVPDIEFIIDNNKRPFLTEKFNTRYVWENAT 892

Db	844	ytvlaoskseeeneemscrrpvtvrtidtnvlnvldmknkraadistfrhevtuyltjemp	903
Qy	893	SGDEVLVASHDRDRELtYtFATVtTNFANvPRtRDFEEVNDLTGtLEvNYHPGEXDLRD	952
Db	904	dftnvvlysidedrdetlytvtvqinayvnprrldtffaldidsqvyuennlmelldt	963
Qy	953	GDEPTHTtPvNFIDNFESDGGKRRNODEVtFVVLVDNDNARvEMtRvDLEtRvSEGA	101
Db	964	rgeqhrhtfnlndtlyfsegdgnrtvntevlrllddenaprdlptlewsisseddq	102
Qy	1013	AGVAVLDE---tYARDBDEPDTNRSvGLDRLTDPRtE-vPRtFMtISTENKtGEL	106
Db	1024	egltlgesdvlyarpdtkedctrsnhyvaimtvcnrdlcvrpllmndspnvcgl	108
Qy	1069	ETADDLtGvNGtYEtFtEADtDHGtPOORSNEtTVLVRtFNtFNvFvRvPORvSvTLRt	112
Db	1084	qtamplrgyvgtydslyadthnlpqstheuyelctelrpnhrppqvfvrsegtllrld	114
Qy	1129	ERATGEGVLtAANEtLEtPvYtADEDGLAGSVtFHVNGEAvOYtJDEtEVGAGENSQ	118
Db	1144	eravnnvslvngpdrtdrtdqaldiddglagvvtfdvlgadadasnfyv - ndgasfgt	120
Qy	1189	LILtQLtPEPQIRQvRtITtRATDGTGtENGtLMTDVTtSVtVvPvRGORvPSvENAtYARtE	124
Db	1202	llltqalpreeketevlratdgttprsysltstltvltvrlpqlrlfqdnlysvafte	126
Qy	1249	GEEGLtRESFELtPQvMEDtKNtCEDDCCDItYtREtDGNEGtLFLvDSSvNTtSLAQLDRE	130
Db	1262	keyglterfslphaedeknklctdcchdllyrtifgvydprfdlrdvntvlfkxeldte	132
Qy	1309	VAtSVtLHtAASNSPDAtGtPLOTtSLvWvNVvNvREANvPREtFEDDLYtAGtSLDStGR	136
Db	1322	tcthavvaasnpctggsprlpsslltvtvltvredprvfegrllytagistcdhntre	138
Qy	1369	LtYtRASHTEDDtTtYtIDRtASQvDSSLEAvVDSAFALHATtGVtSLMQtPRASHNGtF	142
Db	1382	lltvtatshenaqltytiedgsmaavdsleavvdsafhnaqgyvllltlqprasmqmf	144
Qy	1429	EFVDIADtASADtAVtKVtLLtSSORvRtFtDNOLEtEVRKRNtFtAATFtSGFMNTCN	148
Db	1442	etfnvlatpdektdtaevkvylltssqtrvslfndetvesnrdfiaeftsvgfmton	150
Qy	1489	IDQvVtPvFSSDGSvAODDTEvTtEARNtFtIDNPvPOAQEAvRSDtVtLRTtQLMStNSvL	154
Db	1502	ldqylpqtndagvlgaeamevhanffiqdnlpvsadssteelrstqlltvsqylnqrllv	156
Qy	1549	LQDLYtGEDTtPLtGEESQvLAVtALALASAVtGLtCVtLLtALtCFRtPALNtROLQALSMtK	160
Db	1562	lndltvgspsldlgaqygtltvynlaaglsalaeftclllltftvtralnrtleasmek	162
Qy	1609	YGSvDGLNtRAGtL-APtCNtKNHAvEGSvMvNvEAtRvAPDPAISDAGSDSLtGIEMDQ	166
Db	1622	ygsavdsdlntvglaaptnkhallegsnprvneqlkarpdtdaisdtsdcdelljyedsiq	1680

RESULT	4
AAW90182	
ID	AAW90182 standard; Protein; 1528 AA.
XX	
AC	AAW90182;
XX	
DT	05-JUL-1999 (first entry)
XX	
DE	Manduca sexta BT toxin receptor BT-R1.
XX	
KW	BT-R1; receptor; tobacco hornworm; <i>Bacillus thuringiensis</i> ;
KM	BT-toxin; cadherin; insecticide; pesticide.
XX	
OS	Manduca sexta.
XX	
key	Location/Qualifiers
FH	
Peptide	1..21
CT	

CC	FT	Protein	/note= "signal peptide"
CC	FT		22..1528
CC	FT	Domain	/note= "mature protein"
CC	FT		1..1405
CC	FT	Domain	/note= "extracellular domain"
CC	FT		1406..1427
CC	FT	Domain	/note= "transmembrane domain"
CC	FT		1428..1528
CC	FT	Region	/note= "intracellular domain"
CC	FT		72..176
CC	FT	Region	/label= Cad1
CC	FT		177..289
CC	FT	Region	/label= Cad2
CC	FT		290..397
CC	FT	Region	/label= Cad3
CC	FT		398..498
CC	FT	Region	/label= Cad4
CC	FT		499..622
CC	FT	Region	/label= Cad5
CC	FT		623..756
CC	FT	Region	/label= Cad6
CC	FT		757..880
CC	FT	Region	/label= Cad7
CC	FT		881..1004
CC	FT	Region	/label= Cad8
CC	FT		1005..1121
CC	FT	Region	/label= Cad9
CC	FT		1122..1241
CC	FT	Region	/label= Cad10
CC	FT		1242..1352
CC	FT	Region	/label= Cad11
CC	FT		470..500
CC	FT	Peptide	/note= "CNBr peptide"
CC	FT		931..957
CC	FT	Peptide	/note= "CNBr peptide"
CC	FT		1101..1119
CC	FT	Peptide	/note= "CNBr peptide"
CC	FT		1257..1273
CC	FT	Peptide	/note= "CNBr peptide"
CC	XX		
CC	PN	WO9859048-A1.	
CC	XX		
CC	PD	30-DEC-1998.	
CC	XX		
CC	PF	08-JUN-1998;	98WC-US11868.
CC	XX		
CC	PR	20-JUN-1997;	97US-0880042.
CC	PA	(UYWY-) UNIV WYOMING.	
CC	PI	Bulla LA:	
CC	DR	WPI, 1999-105622/09.	
CC	DR	N-PSDB; AAV74277.	
CC	XX		
CC	PT	Newly isolated Bt-toxin receptor proteins	- useful for developing
CC	PT	improved pesticides	
CC	XX		
CC	PS	Claim 1; Fig 2; 62pp; English.	
CC	XX		
CC	CC	The present sequence is Bt-R1, a glycoprotein receptor from the tobacco hornworm which binds to a Bacillus thuringiensis (Bt) toxin. Bt-R1 cDNA (see AAV75277) was obtained from a Manduca sexta midgut tissue cDNA library using degenerate probes based on isolated peptides (see AAW0183-87) of Bt-R1. Amino acid sequence analysis of the deduced protein shows that Bt-R1 is a member of the cadherin family, and has a highly repetitive external domain. The availability of Bt-R1 cDNA permits the retrieval of DNAs encoding homologous receptors in other insects and organisms as well as the position of assays for the cytotoxicity and binding affinity of potential pesticides and the development of methods to manipulate natural and/or introduced homologous receptors and, thus, to destroy target cells, tissues and/or organisms. The cDNA can also	

CC be utilised in the recombinant production of Br-R1 polypeptides
 CC that can be used: as targets for identifying agents that bind to
 CC Br-R1, as targets or bait for identifying and isolating binding
 CC partners and cellular factors that bind to Br-R1, as assay targets
 CC for identifying Br-R1 and other receptor-mediated activity, and as
 CC markers of cells that express a member of the Br-R1 protein family.
 CC They are also used in mediation of cell and tissue destruction,
 CC dissociation, dispersion, cell-to-cell association, and changes in
 CC morphology.

XX Sequence 1528 AA:

Query Match 52.7%; Score 4693; DB 20; Length 1528;
 Best Local Similarity 60.7%; Pred. No 0;
 Matches 925; Conservative 220; Mismatches 357; Indels 22; Gaps 13;

QY 1 MGVERFPVAVLVSLASALANORCYIIAIPR-PETPELPIDIDEGKSMSEQPLIPGPT 59
 1 mavdvriaaflvflapavlaqercymlaprlprpdlpynlfegqtwsgprlpape 60
 QY 60 REEVCMEF---LPDQNIQVIYMEETEGDYIAKLNYQGSNTPLVSI--MSGOPRAQLG 114
 61 rddicmdahvitanlglqylmdeeideftlaalngnpsrpflelplfsgsynl-lm 119
 QY 115 PEKQNEADQMSLVITQRDYEATQMSVVSIOVEGSAVLVALEIVINDNPRILQ 174
 120 pvlrr---vdngsaahhaharqnyelpgmqymftrvdqgsilvagvsalvnlidnapilq 177
 QY 175 VSAACVIRPEHGEARLDCVYQVSDRDEISTRFMTFRVDSRRADESIFYWVGSDP 234
 178 nfeprcypelgeplrectyqvdadgristefmtfidsvr-gdeeflyertnlpngw 236
 QY 235 FNMKMTVGINSPLNEFTTOLHFSVTSASDSLPNNHVTYMAVOVENESRPPKWEIFSVQ 294
 237 mwimltgvtstnftvspilhsftaldslprlthvtmnygvansrprvlelfavq 296
 QY 295 QPDEKTOQSLSRAIDODTGINRAINTLLRDADDFESLEVI---BDGALHAYTEIDRD 351
 297 qfeeksyqnlvraiddetelmpnlrlltneedtfisleaipggsgsvflvpslrd 356
 QY 352 KLERELFNLTIVAYKSTDSAFATEAHFIIVNDVQRPPLHKEYSIDMEETPMTLNF 411
 357 tlgrevpltiavkykeaeafststnvlvtidndqrpelkaylameetpltnf 416
 QY 412 NEERGFHDRLGENAQYVLELVFPPGAASAFYIAPGSGYOROTFTMGTINTMIDYED 471
 417 dkefgfndkldgnaqytrlesvdpgraaefyiapevgyqrqtfimglshmsldyev 476
 QY 472 VIRQNIITIKKAVDMNNAHVGEALYVNLINNDLPFEESYSASFETVAGAPVA 531
 477 pqlgsitlvraiddndtrgvahvldilnmdegpllehavqvrltldetsegffva 536
 QY 532 TVLALDRIDIDVVVHSMGNAVYULFIDESTGELFVSMDDAPFYHQNTFVQVRADDTL 591
 537 kavahtddidgvvehltlignavnfltdkltdgtrivsandsfnyheselfvyratdcl 596
 QY 592 GDGPHTVTQOLVIELEDVNNTPPLRLPRSTPSVEENVEGEYISREITATPDPTSAYL 651
 597 ge-pfhatsqlvrlndnlnrptlrlprspqveenypdghvltqelratpdtlacl 655
 QY 652 WPEIDMSTWATQGRRTNTEYVGCIVFTIYPTGKNRSALGRVLVDIRNVTIDEE 711
 656 rfeinwtslatskqgrqanqetfrnceletlfrlnnrglaigrvareltlnhvtdyge 715
 QY 712 EPEMLVTVVRDLNMYIGDYDEATFTIIFIDMNDAPFFAGTTLQVRFVRELAASGT 771
 716 efevlstlvtrvrdlnnygdqydesmlltclimdnnapvwvgeltleqnftrvemsagql 775
 QY 772 LIGSVLATFDIGPLYNQVRYTIOPRNPTPEGVLKIDFTTGQIEVDANEAIDAPMRFYL 831
 776 vvgsvradddgpllyngvrytlfrpredtkdlimidfqltglsivntsgaldadtprfhl 835

QY 832 YTVVIASDECSLENRRECPDPSNVFEPVDIEIETIDTNKVPDEPLEKENTVYVWENA 891
 836 yltvvaadrcstcpadcpdpdytweegnltlhtcdtmkvpqaetcktdltvyllyena 895
 QY 892 TSGDEVOALYSHDRDELHYTRYTMNFANRPLRDLFEFVDDTGRLEVHYF--GDEK 948
 896 thldevvtliasadldrdelqyhtvsyalyanvnpnlmffesvretqlvydydtqsgsev 955
 QY 949 LDRDGEPTHTFVNFIDNFSDDGRRNDEVEIFVVLVDVNDNAPEMLPDELFDVS 1008
 956 ldrdgepthtfifnldnfmgegnnqndevlilldvndnapelpppselwtls 1015
 QY 1009 EGAVAGVRLPEIYAPDRDPTDNRSGVGIIDLTFTTDRIDVPPDLFTWISLENTGEL 1068
 1016 enlkgyvrlphlfapdrdpdcnsvgyellnls-terdlevpelfrmqlanvtgel 1074
 QY 1069 ETAMDLRGWYGEIETLEADHGYPOORSNETYLYIRYFNHFPVFPPODPDVIRLSR 1128
 1075 etamdltkgywtyalhhtrdthgipqsmmeteyellhprfnypetfvrftndavrlar 1134
 QY 1129 ERATEGVLAFAANERLEPIYATDEGLAAGSVTFHVQNEEAVOYFDITTEGAGENSQ 1188
 1135 eravngvlatvngelrlexlratcpdglhagvvtfgvvgdeesqrfyv--vndgenlgs 1192
 QY 1189 LILROLFPEQIRQFRITIRATDGTGEPGLMTDVPFSVFPVPGSPVSENAATVAFPE 1248
 1193 lrlilgvpelrefrlrltalatdgtdpgrlstmtctlrvtvfpqgeprlasehavaflle 1252
 QY 1249 GEGELRESEPLEQABDLKNHLCEDDCODIYRFDGNNEGLFVLDDSSNVISLAQELDE 1308
 1253 ksagmesnqlplaqdtkhnlceddchsllyrrlidngsnghfgldpvrnrlflkkelire 1312
 QY 1309 VATSYTLHIAASNSPRTATGPILOTSLVTVVNVREANPRPIEQDLYTIGISTIGRE 1368
 1313 qsashtlqvaaasnpsdg-9lpasallvtvtrreadpvrfrllytagistadsigre 1371
 QY 1369 LITVRASHTEPDRTITVTIRASQWOLDSLEAVRDSAFALHATGVLISLMOPASMHGMF 1428
 1372 ltrlhacqgsaalcyaldtdlnvpsleavqsaftvlnaagvrltlnlqpratahmglt 1431
 QY 1429 EFDVIATDPAASAITARVVVYLLISSONRYTFLEFNDQLETVQONRNFIAATFSGFMNTGN 1488
 1432 kfevatctdtagadrdvrcvyvssqnrtyfvtvnlqgvendrdtadtfaagfmntcn 1491
 QY 1489 IDQVPPSDS--SGVAQDDTTEVRA 1511
 1492 idqvvpandpvtgvalehstqmaa 1515
 Db

RESULT 5
 AAB20500 standard; Protein; 1734 AA.
 AC AAB20500;
 DT 09-JUL-2001 (first entry)
 XX
 DE Fall armyworm Bt toxin receptor.
 XX
 KW Fall armyworm; FAM: Bacillus thuringiensis; Bt toxin;
 KW receptor; crystal protein; CryIA; biological control; insecticide;
 KW crop protection.
 XX
 OS Spodoptera frugiperda.
 XX
 FH Key location/Qualifiers
 FT Peptide 1..22
 FT Protein /label-Signal-peptide
 FT 27..1734
 FT Binding-site /label-Mature_protein
 FT 1317..1486
 FT Domain /label-CryIA_binding_site
 FT 1594..1612

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2002, 17:08:22 ; Search time 70.14 Seconds
(without alignments)
1864.723 Million cell updates/sec

Title: US-09-715-909-2
Perfect score: 8909
Sequence: 1 MGVERFPVLLVSLASAL.....THGNNFGPKSPYLPDPHPK 1717

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Sequenced: 219241 seqs, 76174552 residues

TC: number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

1: PIR_68:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5607.5	62.9%	1715	JE0128	Bombyx mori receptor
2	602	6.8	5147	JE0128	cadherin-related t
3	560	6.3	4351	T00252	MEGFL protein - ra
4	516.5	5.8	3097	T00021	DN-cadherin - fru
5	498.5	5.6	2809	T30213	G-cadherin - sea u
6	492	5.5	1311	T33757	hypothetical prote
7	484.5	5.4	1439	T27110	hypothetical prote
8	452	5.1	3034	T14119	seven-pass transme
9	434.5	4.9	4307	T15276	hypothetical prote
10	405	4.5	2163	T120721	hypothetical prote
11	403	4.5	2610	T20968	hypothetical prote
12	374	4.2	3343	S44887	hypothetical prote
13	369.5	4.1	1965	S75200	ZK112.7 protein -
14	318	3.6	1069	T00043	fat protein - Syne
15	311	3.5	1069	T00043	BH-protocadherin-a
16	308.5	3.5	1072	T31066	vascular cadherin-
17	293	3.3	1200	T00041	BH-protocadherin P
18	291.5	3.3	882	T00042	BH-protocadherin P
19	290	3.3	1544	T29482	cadherin 1 precurs
20	289.5	3.2	906	T29482	hypothetical prote
21	283.5	3.2	889	T09055	N-cadherin precurs
22	279	3.1	785	T09055	protocadherin 68 -
23	274	3.1	906	T09055	cadherin-7 - chick
24	273	3.1	796	A38992	cadherin 2 precurs
25	273	3.1	796	I48277	cadherin 11 precur
26	271.5	3.0	790	I48277	cadherin-11 - mous
27	270.5	3.0	877	I7016	cadherin-6 - human
28	269	3.0	701	I7016	N-cadherin precurs
29	269	3.0	796	T17243	hypothetical prote
30	269	3.0	796	A53584	OB-cadherin precu

30	267.5	3.0	790	2	I50178	cadherin-6B - chic
31	267	3.0	789	2	I52701	K-cadherin - rat
32	266	3.0	2205	2	T08615	aggregation factor
33	264.5	3.0	794	2	I59372	cadherin 12 - huma
34	264.5	3.0	796	2	I49556	cadherin-11 - huma
35	259.5	2.9	790	2	G02678	cadherin-14 - huma
36	258	2.9	790	2	I51638	F-cadherin - Afric
37	257	2.9	832	2	S55396	Li-cadherin - huma
38	251.5	2.8	871	2	S47518	cadherin - African
39	249.5	2.8	827	2	A53954	Li-cadherin precu
40	248.5	2.8	1329	2	T33136	hypothetical prote
41	243.5	2.7	906	1	IJXLC2	N-cadherin 2 precu
42	241	2.7	793	2	D38992	cadherin 8 - human
43	238.5	2.7	783	2	I50116	N-cadherin precurs
44	235.5	2.7	912	1	I50116	N-cadherin precurs
45	235	2.6	784	1	IJHUC5	cadherin 5 precurs

ALIGNMENTS

Query Match	Best Local Similarity	Score	Pred. No.	DB 2	Length	Indels	Gaps
Matches 1095;	Conserved 224;	Mismatches 377;	Indels 29;	Gaps 12;			
QY	1	MGVERFPVLLVSLASALANQRCYIIAIPRPETPELPIDEGSKSWSEQPLIPGPR	60				
DB	1	MGVDVRLATLLIYAEYVLAQERGFVAIPRPDPLEIDFEGQWQSRPLIPADR	60				
QY	61	EYVCMENF--LPDMIOYIYMEEEIEGVYIYAKLNYQGSNTPVL--STMSCOPRAQLCP	115				
DB	61	EYVCMGVHAMPYTGTOIYMEEEIEGVYIYAKLNYRPNVYIEPARLSGSFNLVYV	120				
QY	116	EPFQNMADQMSLVITRDQYFATQMSYVSIOVGEQSAVLVALEIYINIDNPPIIV	175				
DB	121	IRRIIPDSNEMHILITRDQYFATQMSYVSIOVGEQSAVLVALEIYINIDNPPIIV	180				
QY	176	VSACVPIPEHGEARLDCVYQVSDRDELSTRPMTFRVDSRADESIFVGEYDPSDMF	235				
DB	181	LEPCQVDELGEARLDCVYQVSDRDELSTRPMTFRVDSRADESIFVGEYDPSDMF	239				
QY	236	NNKATVGINSPLEFETQIHLFSVTSASDSLPNNHYVTMMQVYENSRPRVVEIFSVQO	295				
DB	240	RMTVMVGINSPLEFETQIHLFSVTSASDSLPNNHYVTMMQVYENSRPRVVEIFSVQO	299				
QY	296	EPEKTNQSLRAIDGDSINRAIVTILRDADDFESLEVE--DGAIIHVFTEIDRK	352				
DB	300	FPEKTRQSPVRAIDGDSINRAIVTILRDADDFESLEVE--DGAIIHVFTEIDRK	359				
QY	353	LRRELFLNLTIVAKSSTDSAFATEAHLFIIVNDVNDORPEPLHKEYSIDIMEETPTLNFN	412				
DB	360	LRREVFQSLTIYAKSSTDSAFATEAHLFIIVNDVNDORPEPLHKEYSIDIMEETPTLNFN	419				

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 15, 2002, 17:19:57 ; Search time 64.19 Seconds
(without alignments)
980.737 Million cell updates/sec

Title: US-09-715-909-2

Sequence: 1 MGVRRFPVAVLVSLASAL.....THGNNGFKSTPLPQPHK 1717

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 3664827 residues

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	619	6.9	4590	1	FATH_HUMAN
2	594	6.7	5147	1	FAT_DROME
3	516.5	5.8	3097	1	CADN_DROME
4	374	4.2	3343	1	YOG7_CAEL
5	339.5	3.8	789	1	CAD9_HUMAN
6	337.5	3.8	949	1	CDAB_HUMAN
7	328	3.7	948	1	CDAB_HUMAN
8	325.5	3.5	950	1	CDAB_HUMAN
9	314.5	3.5	1507	1	CADE_DROME
10	312.5	3.5	950	1	CADE_HUMAN
11	311.5	3.5	798	1	CBDE_HUMAN
12	310	3.5	948	1	CDAC_HUMAN
13	308.5	3.5	941	1	CDAC_HUMAN
14	307	3.4	947	1	CDAC_HUMAN
15	307	3.4	1007	1	CDAC_HUMAN
16	303.5	3.4	950	1	CDAC_HUMAN
17	303	3.4	788	1	CADA_HUMAN
18	301	3.4	818	1	CDBI_HUMAN
19	299	3.4	818	1	CDBI_HUMAN
20	297.5	3.3	930	1	CDAB_HUMAN
21	296	3.3	936	1	CDAB_HUMAN
22	295.5	3.3	797	1	CDAB_HUMAN
23	293	3.3	950	1	CDAB_HUMAN
24	291.5	3.3	797	1	CDAB_HUMAN
25	291.5	3.3	882	1	CDAB_HUMAN
26	290	3.3	776	1	CDAB_HUMAN
27	289.5	3.2	906	1	CAD2_MOUSE
28	289.5	3.2	930	1	CDAB_HUMAN
29	289	3.2	795	1	CDAB_HUMAN
30	288.5	3.2	801	1	CDAB_HUMAN
31	288	3.2	797	1	CDAB_HUMAN
32	286.5	3.2	795	1	CDAB_HUMAN
33	281	3.2	795	1	CDAB_HUMAN

34	278.5	3.1	906	1	CAD2_RAT	O9y13 rattus norv
35	277	3.1	800	1	CDAB_HUMAN	O9un67 homo sapien
36	275	3.1	793	1	CDAB_HUMAN	O9y52 homo sapien
37	274	3.1	906	1	CAD2_HUMAN	P19022 homo sapien
38	273	3.1	796	1	CAD6_HUMAN	P55287 homo sapien
39	271.5	3.0	790	1	CAD6_HUMAN	P55285 homo sapien
40	270.5	3.0	877	1	CAD2_BOVIN	P19534 bos taurus
41	270.5	3.0	937	1	CDAB_HUMAN	O9un72 homo sapien
42	269.5	3.0	792	1	CAD6_CHICK	O93319 gallus gall
43	269	3.0	796	1	CAD6_MOUSE	P53288 mus musculu
44	267.5	3.0	790	1	CAD6_CHICK	O90762 gallus gall
45	267	3.0	789	1	CAD6_RAT	P55280 rattus norv

ALIGNMENTS

RESULT 1	FATH_HUMAN	STANDARD:	PRT:	4590 AA.
AC	O14517:			
DT	20-AUG-2001 (Rel. 40, Created)			
DT	20-AUG-2001 (Rel. 40, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	CADHERIN-RELATED TUMOR SUPPRESSOR HOMOLOG PRECURSOR (FAT PROTEIN			
DE	HOMOLOG).			
GN	FAT.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lymphocytes;			
RX	MEDLINE=96163873; PubMed=8586420;			
RA	Dunne J., Hanby A.M., Poulson R., Jones T.A., Sheer D., Chin W.G.,			
RA	Da S.M., Zhao Q., Beverley P.C.L., Owen M.J.;			
RT	"Molecular cloning and tissue expression of FAT, the human homologue			
RT	of the Drosophila fat gene that is located on chromosome 4q34-q35 and			
RT	encodes a putative adhesion molecule.";			
RT	Genomics 30:207-223(1995).			
CC	- FUNCTION: COULD FUNCTION AS A CELL-ADHESION PROTEIN.			
CC	- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).			
CC	- TISSUE SPECIFICITY: EXPRESSED IN MANY EPITHELIAL AND SOME			
CC	ENDOTHELIAL AND SMOOTH MUSCLE CELLS.			
CC	- SIMILARITY: CONTAINS 34 CADHERIN DOMAINS.			
CC	- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.			
CC	- SIMILARITY: CONTAINS 1 LAMININ G-LIKE DOMAIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@sib-sib.ch).			
CC	-----			
DR	EMBL: X87241; CAA60685.1; -			
DR	HSSP: P01132; IEPH.			
DR	MIM: 600976; -			
DR	InterPro: IPR000152; Asx_hydroxyl.			
DR	InterPro: IPR002126; Cadherin.			
DR	InterPro: IPR000561; EGF-like.			
DR	InterPro: IPR000742; EGF-2.			
DR	InterPro: IPR001881; EGF-Ca.			
DR	InterPro: IPR001791; laminin_G.			
DR	pfam: PF00028; cadherin_33.			
DR	pfam: PF00008; EGF_5.			
DR	pfam: PF00054; laminin_G_1.			
DR	PRINTS: PR00205; CADHERIN.			
DR	SMART: SM00112; CA; 32.			
DR	SMART: SM00179; EGF_CA; 1.			
DR	SMART: SM00001; EGF_like; 5.			

